1971

Some statistical inferences for the bivariate exponential distribution

Bruce Mohr Bemis

SOME STATISTICAL INFERENCES FOR THE BIVARIATE

EXPONENTIAL DISTRIBUTION

by

BRUCE MOHR BEMIS, 1936 -

A DISSERTATION

Presented to the Faculty of the Graduate School of the

UNIVERSITY OF MISSOURI - ROLLA

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

MATHEMATICS

1971

Lee J. Bain
Advisor

Troy L. Hicks

James J. Higgins

Caslav V. Stanojevic

Donald H. Howley

Billy E. Gillett

ABSTRACT

The bivariate exponential distribution is neither absolutely continuous nor discrete due to the property that there is a positive probability that the two random variables may be equal. Basic properties of the distribution are presented as well as methods of parameter estimation including maximum likelihood. The distribution is shown to satisfy the usual regularity conditions in spite of its possession of a singularity. The maximum likelihood estimates are asymptotically efficient. Two other methods of estimation are compared with the maximum likelihood method in terms of efficiency.

Tests of the hypothesis that two random variables each have independent exponential distributions versus the alternative hypothesis that the variables follow a bivariate exponential distribution with positive correlation are considered in detail. The estimation of the reliability of a simple two component series system or a parallel system in which the components have life times which follow the bivariate exponential distribution is considered. The errors made when assuming erroneously that the two random variables are independent, each with exponential distributions, when in fact they follow the bivariate exponential distribution, are illustrated.

ACKNOWLEDGEMENT

The author expresses his appreciation to Dr. Lee J. Bain of the Department of Mathematics for his help in the selection of this thesis subject and for his assistance and encouragement during its preparation. Also appreciation is expressed to Dr. James Higgins for consenting to serve as second reader.

The author is indebted to the National Science Foundation for the financial support they provided. The author is indebted to Westminster College (Utah) for the sabbatical leave granted for this time of study.

Deepest gratitude is expressed to my wife, Joy, for her diligent typing of this thesis and especially for her patience and encouragement during these years of study.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

I. INTRODUCTION

A multivariate exponential distribution which allows for dependency among the variables has recently been introduced in the literature [1]*. It appears that this distribution will be an important model for reliability studies as well as being an interesting distribution from a mathematical point of view. The object of this thesis is to consider certain statistical inferences about the bivariate exponential distribution. It is felt that an understanding of the bivariate exponential distribution is fundamental to the analysis of the multivariate exponential distribution. The random variables X and Y are distributed according to the bivariate exponential distribution, $(X,Y) \sim BVE(\lambda_1, \lambda_2, \lambda_3)$, if $P[X>x, Y>y] = \exp\{-\lambda_1 x - \lambda_2 y - \lambda_3 \max(x,y)\}$.

The bivariate exponential distribution has both an absolutely continuous and a continuous singular part due to the property that $P[X=Y]>0$ and yet the Lebesgue measure of the set $X = Y$ is zero. This situation causes some problems in maximum likelihood estimation and hypothesis testing.

Some basic properties of the distribution are given in Chapter III. The maximum likelihood equations for the estimates of the parameters are derived and given in Chapter IV. The regularity conditions for the BVED are verified so that the conclusion may be drawn that the maximum likelihood estimates are asymptotically efficient. The information matrix based on a sample of size n for the distribution is obtained.

*All numbers [a] refer to the bibliography while the numbers (a) refer to equations.

In Chapter V another method of estimation of the parameters is given which is based on the idea of the method of moments. The estimators by this method are easier to compute than by the method of maximum likelihood and yet the efficiency is close to that of the method of maximum likelihood. An asymptotic result is given for the covariance matrix of these estimators. It is noted that the proportion of sample points such that X=Y does a better job of estimating the correlation of X and Y than the usual sample correlation coefficient. A comparison of the efficiencies of three methods of estimation is given based on simulations from the bivariate exponential distribution and asymptotic considerations.

Chapter VI deals with the problem of testing the hypothesis that the correlation between X and Y is zero against the alternative that the correlation is positive when $(X,Y) \sim BVE(\lambda_1, \lambda_2, \lambda_3)$. That is, procedures are given to determine whether a sample comes from a joint distribution which is the product of marginal exponential distributions, or from the bivariate exponential distribution with $\rho = cor(X,Y) > 0$. A simple test based on the number of sample points such that X=Y is given first. This test is useful when none of the parameters is known and the power of the test serves as a lower bound for other tests considered. The power of this test is tabled for various values of ρ and sample sizes n. It is indicated how this test may be generalized to the multivariate exponential distribution.

When all parameters but ρ are known then a test of the simple hypothesis $H: \rho = 0$ against the simple alternative $K: \rho = \rho_o > 0$ may be based on the method of the Neyman-Pearson lemma. This test is derived and critical values and power tables are given. When the marginal parameters $\gamma_1=\lambda_1+\lambda_3$, $\gamma_2=\lambda_2+\lambda_3$ are known and equal, a uniformly most powerful test exists for H : $\rho=0$ against K : $\rho>0$. When $\gamma_1=\gamma_2$ but the common value is unknown, a test procedure is given which has power close to the power of the test mentioned above. For each of these tests considered, as n gets large or as ρ gets large the power approaches the power of the test given in the preceding paragraph. A confidence region is given for the sum of the parameters $\lambda_1+\lambda_2+\lambda_3$. Then a joint confidence region is obtained for the correlation ρ between X and Y and the sum of the marginal parameters of the marginal exponential distributions. From this a conservative confidence interval may be developed for the sum of the marginal parameters.

In Chapter VII the problem of estimating the reliability of a simple two component series system or a two component parallel system is considered where the lifetimes of the two components X and Y, are assumed to follow the bivariate exponential distribution. Point estimates of both the series reliability and the parallel reliability are given and a lower confidence limit for the series reliability is developed.

If when estimating reliability it is assumed that X and Y are independent when in fact they are dependent an error will be introduced. A comparison of this discrepancy is given for various sample sizes and values of the correlation between X and Y based on simulations. It is found that the series reliability estimate is conservative in that it underestimates the true value when assuming wrongly that X and Y are independent whereas the parallel reliability estimate is misleading in that it overestimates the true value when assuming wrongly that X and Y are independent. A further example of the introduced error in assuming wrongly that X and Y are independent when $(X,Y) \sim BVE$ is given in terms of a tolerance region.

II. REVIEW OF THE LITERATURE

In 1966 Marshall and Olkin [1] gave meaningful derivations of a multivariate exponential distribution based on shock models and the requirement that residual life is independent of age. The distribution obtained has the interesting feature of having both an absolutely continuous part as well as a continuous singular part. Meaningful examples of this type of distribution have been missing from the literature and many standard textbooks on statistics seem to dismiss from consideration the case where the distribution is neither discrete nor absolutely continuous. The bivariate exponential distribution is given by $\overline{F}(x,y)=P[X>x,\ Y>y]=\exp\{-\lambda_1 x-\lambda_2 y-\lambda_3 \max(x,y)\}$ and will be denoted by $BVE(\lambda_1,\ \lambda_2,\ \lambda_3)$. The moment generating function for the bivariate exponential distribution is obtained in [1] as well as the important property that if X and Y are dependent but (X,Y) is distributed according to the bivariate exponential distribution then min(X,Y) has an exponential distribution. It was shown that $(X,Y)\sim BVE$ if and only if there exist independent exponential random variables U, V, W such that X=min(U,W) and Y=min(V,W). An initial discussion is given concerning the discrepancy in assuming that the random variables are independent with exponential marginals when in fact they are dependent and follow the bivariate exponential distribution. The maximum discrepancy between $\overline{F}(x,y)$ and $\overline{F}_1(x)\overline{F}_2(y)$ in terms of the correlation between X and Y is given as $[\rho^{\rho}/(1+\rho)^{1+\rho}]^{1/\rho}$ where $\overline{F}(x,y)$ denotes $P[X>x,\ Y>y]$ when $(X,Y)\sim BVE$.

In 1968 Arnold [2] dealt with the problem of parameter estimation for the multivariate exponential distribution. He gave a method of estimation which is unbiased and consistent for the parameters. A measure of efficiency of the estimation procedure was given using the criterion of the mean square of the vector of estimates, that is, the expected squared distance of the vector of estimates from the vector of parameters.

For the bivariate case the information matrix is given in [2] and it is stated that the infinum of the asymptotic relative efficiency of this method of estimation is 1/2 in the limiting case when $\lambda_1 = \lambda_2$ is large relative to λ_3. Also if the three parameters are equal then the asymptotic relative efficiency is approximately 81/105 whereas when one of the parameters is much larger than the other two the asymptotic relative efficiency is close to 1.

Though only these two articles deal directly with the bivariate exponential distribution, this new probability model has been found in the textbook Applied Probability by Thompson [3] as an example of a reliability model for dependent responses. In the article entitled "A Reliability Bound for Systems of Maintained, Interdependent Components" [4] the concept of associated random variables which was discussed in [5] is mentioned with the multivariate exponential distributiontion as a prime example. Also the concept of a joint performance process being associated in time is mentioned with the Marshall-Olkin multivariate exponential distribution an example. Finally in this same article [4] it is mentioned that the case of a "2 out of 3" system, when the components have lives distributed according to the trivariate exponential distribution, gives an example to show that the minimal cut lower bound is a poor approximation to system reliability due to the fact that there is a positive probability that two or more components may fail simultaneously.

One further place where the multivariate exponential distribution is mentioned as an example is in the article [6] in which a definition is given for a multivariate distribution function to have an increasing hazard rate.

In the abstracts of the April 1969 issue of the Annals of Mathematical Statistics [7] Maik indicates that he has found the joint asymptotic distribution of the i^{th} order statistic from the bivariate exponential distribution. In the abstracts of the April 1970 issue of the same journal [8] Maik indicates that he has found a set of complete sufficient statistics for the parameters of the multivariate exponential distribution and hence the unique minimum variance unbiased estimates of these parameters within the framework of observing only the minimum of the responses and which variable yielded the minimum.

III. SOME PROPERTIES OF THE BIVARIATE EXPONENTIAL DISTRIBUTION

A. Introduction

In [1] basic properties of the multivariate exponential distribution are given with specific reference to the bivariate exponential distribution. The purpose of this chapter is to restate some of those properties given in [1] and to state other properties which may be used in later chapters. In particular the density function for the bivariate exponential distribution which does not exist with respect to the usual two dimensional Lebesgue measure will be given under the condition that the two random variables lie in certain regions. Also the manner of integration with respect to the bivariate exponential distribution will be considered with an example to illustrate the method. The central moments of the bivariate exponential distribution are tabulated through the fourth order. Finally a comparison will be made between the distribution of $\overline{X}$ and $\overline{Y}$ based on a sample of size n from the bivariate exponential distribution with an appropriate bivariate normal distribution and a certain bivariate gamma distribution given by Kibble in [9].

B. Basic Properties of the Bivariate Exponential Distribution

Let X and Y be continuous random variables each of which take values in (o,∞). X and Y are distributed according to the bivariate exponential distribution (BVED) with parameters $\lambda_1,\lambda_2,\lambda_3$ if $P[X>x,Y>y]=\exp\{-\lambda_1 x-\lambda_2 y-\lambda_3 \max(x,y)\}$[1] where λ_1 and λ_2 are positive and λ_3 is nonnegative. For convenience of notation this is indicated by $(X,Y)\sim BVE(\lambda_1,\lambda_2,\lambda_3)$. In case $\lambda_3=0$ then $(X,Y)\sim BVE(\lambda_1,\lambda_2,0)$ is equivalent to saying that X and Y are independent, each with exponential distributions with parameters λ_1 and λ_2, respectively. Letting $\overline{F}(x,y)=P[X>x, Y>y]$ it is clear that the marginal distributions of X and Y are respectively, $\overline{F}_1(x)=\overline{F}(x,0)=\exp\{-\lambda_1 x-\lambda_3 x\}$ and $\overline{F}_2(y)=\overline{F}(0,y)=\exp\{-\lambda_2 y-\lambda_3 y\}$. If $\gamma_1=\lambda_1+\lambda_3$ and $\gamma_2=\lambda_2+\lambda_3$ then X has a marginal exponential distribution with parameter γ_1 and Y has a marginal exponential distribution with parameter γ_2.

It is pointed out in [1] that the BVED has both an absolutely continuous and a singular part due to the fact that $P[X=Y]=\lambda_3/\lambda$ where $\lambda=\lambda_1+\lambda_2+\lambda_3$, while the Lebesgue measure of the set $A=\{(x,y)\mid x=y>0\}$ is zero. That is, this distribution is continuous but not absolutely continuous with respect to ordinary Lebesgue measure on $(o,\infty)x(o,\infty)$. Let $\overline{F}(x,y)=G_1(x,y)+G_2(x,y)$ where $G_1(x,y)=\exp\{-\lambda_1 x-\lambda_2 y-\lambda_3 \max(x,y)\}-\lambda_3 \exp\{-\lambda\max(x,y)\}/\lambda$ and $G_2(x,y)=\lambda_3\exp\{-\lambda\max(x,y)\}/\lambda$. G_1 is absolutely continuous with respect to the usual two dimensional Lebesgue measure and G_2 is continuous singular concentrated on $x=y>0$. The mixed partial of G_1 with respect to x and y is

$$\frac{\partial^2 G_1}{\partial x\partial y}=\begin{cases}\lambda_1\gamma_2\overline{F}(x,y) & \text{if } x<y\\ \\ \lambda_2\gamma_1\overline{F}(x,y) & \text{if } x>y\end{cases}$$

This may be interpreted as the density of the bivariate exponential distribution over the regions $R1=\{(x,y)\mid 0<x<y\}$ and $R2=\{(x,y)\mid 0<y<x\}$. The ordinary derivative of $G_2(x,x)$ with respect to x is $\lambda_3\overline{F}(x,x)$; this may be interpreted as the density of the bivariate exponential distribution over $R3=\{(x,y)\mid 0<x=y\}$.

The above is motivated by the following considerations.

$$1=\int_0^\infty\!\!\int_0^\infty d\overline{F}(x,y)=\int_0^\infty\!\!\int_0^\infty d(G_1+G_2)=\iint_{R1}d(G_1+G_2)+\iint_{R2}d(G_1+G_2)+\iint_{R3}d(G_1+G_2)$$

Now $\int\int d(G_1+G_2)=\int\int dG_1+\int\int dG_2$. Since G_2 is concentrated on R3 the contributions of $\int\int dG_2$ over R1 and R2 are zero. Since G_1 is absolutely continuous, the contribution of $\int\int dG_1$ over R3 is zero. Hence

$$1=\int_0^\infty\int_0^\infty d\overline{F}(x,y)=\int\int_{R1} dG_1+\int\int_{R2} dG_1+\int\int_{R3} dG_2.$$

But each of the integrals on the right side in the above equation may be evaluated by ordinary methods.

$$\int\int_{R1} dG_1=\int\int_{R1}\lambda_1\gamma_2\overline{F}(x,y)\,dxdy=\lambda_1/\lambda$$

$$\int\int_{R2} dG_1=\int\int_{R2}\lambda_2\gamma_1\overline{F}(x,y)\,dxdy=\lambda_2/\lambda$$

$$\int\int_{R3} dG_2=\int_0^\infty\lambda_3\overline{F}(x,x)\,dx=\lambda_3/\lambda \qquad \text{(a)}$$

To indicate why the first equality in equation (a) above holds consider the following. Let $S=(0,M)\times(0,M)$ where M is a positive constant and let $T=R3\cap S$. Let P be a partition of S into rectangles $\{R_i\},(i=1,\ldots n)$. Let the rectangles in the partition P be divided into two disjoint sets P_1 and P_2 where P_1 is the set of rectangles in P which have no interior points common to T and P_2 is the set of rectangles in P which have interior points common to T. For example if $P=\{R_i\},(i=1,\ldots,15)$ where the R_i are illustrated in Figure 1 below, then P_1 is the set of rectangles numbered 1, 2, 3, 4, 7, 8, 9, 11, 14 and P_2 is the set of rectangles numbered 5, 6, 10, 12, 13, 15.

Let $G_2|R=G_2(a,b)+G_2(c,d)-G_2(a,d)-G_2(c,b)$ where $R=\{(x,y)|a<x<c, b<y<d\}$. Then $\int\int_{R3} dG_2=\sum G_2|R_i$.* For R_i in set P_1 it is clear from the definition of G_2 that $G_2|R_i=0$ since R_i above T implies $G_2(a,b)=G_2(c,b)$ and $G_2(a,d)=G_2(c,d)$ whereas R_i below T implies $G_2(a,b)=G_2(a,d)$ and

*The index of $\sum$ will always be from 1 to n unless indicated otherwise.

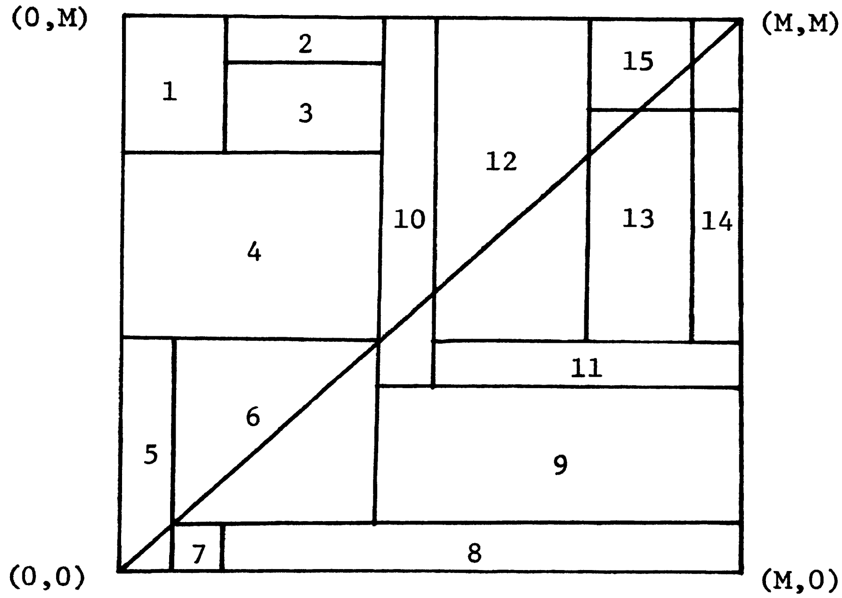

Figure 1. A Partition of the Set S

$G_2(c,b)=G_2(c,d)$. Hence only rectangles in P_2 need to be considered. Each rectangle in P_2 may be partitioned further to include one square with diagonal common to T and other rectangles which are back in class P_1. Hence only squares with diagonal common to T need to be considered.

Let $0=x_o<x_1<x_2\ldots<x_n=M$ induce a partition of n squares $\{R_i\}$, $(i=1,\ldots,n)$ where $R_i=\{(x,y)|x_{i-1}<x<x_i,\ x_{i-1}<y<x_i\}$. So

$$\iint_{R3} dG_2(x,y) = \sum G_2|R_i = \lambda_3\sum[\exp\{-\lambda x_{i-1}\}-\exp\{-\lambda x_i\}]/\lambda.$$

Let $H(x)=\lambda_3\exp\{-\lambda x\}/\lambda$, and $H|I=H(b)-H(a)$ where $I=[a,b]$. Then $\int_o^M dH(x) = \sum H|I_i$ where $I_i=[x_{i-1},x_i]$ and $0=x_o<x_1<x_2<\ldots<x_n=M$ is a partition of $[0,M]$. Now $H(x)=G_2(x,x)$ so $\iint_{R3\cap S} dG_2(x,y) = \int_o^M dG_2(x,x)$. Taking the limit as $M\to\infty$ it is clear that $\iint_{R3} dG_2(x,y) = \int_o^\infty dG_2(x,x)$.

The function $f(x,y)$ given below by (1) will be referred to as the density function of (X,Y) when $(X,Y) \sim BVE(\lambda_1,\lambda_2,\lambda_3)$. It should be noted that the density for (X,Y) in R1 or R2 is with respect to two-dimensional Lebesgue measure whereas for (X,Y) in R3 the density is with respect to one-dimensional Lebesgue measure.

$$f(x,y) = \begin{cases} \lambda_1\gamma_2\overline{F}(x,y) \text{ if } (x,y)\varepsilon R1 \\ \lambda_2\gamma_1\overline{F}(x,y) \text{ if } (x,y)\varepsilon R2 \\ \lambda_3\overline{F}(x,x) \text{ if } (x,y)\varepsilon R3 \end{cases} \tag{1}$$

The conditional distributions are given below. Though the marginal distributions are absolutely continuous with respect to the one dimensional Lebesgue measure the conditional distributions do not share this property for they illustrate one dimensional distributions which are neither absolutely continuous nor discrete. Note that each has a jump at $x=y$ indicating a point mass there as a reflection of the situation that $P[X=Y]$ is not zero.

Let I_x be the event that X lies in $(x',x]$ and I_y be the event that Y lies in $(y',y]$. Then $F(x,y|I_x)=P[X<x,Y<y \cap I_x]/P[I_x]$ and $F(x,y|I_y)=P[X<x,Y<y \cap I_y]/P[I_y]$. $F(y|x)$ denotes $\lim_{x'\to x-} F(x,y|I_x)$ and $F(x|y)$ denotes $\lim_{y'\to y-} F(x,y|I_y)$. These definitions are given in Wilks on page 60 [10]. In $F(y|x)$ it is assumed that $x>0$ and in $F(x|y)$ it is assumed that $y>0$.

$$F(y|x) = \begin{cases} 0 \text{ if } y<0 \\ 1-\exp\{-\lambda_2 y\} \text{ if } 0<y<x \\ 1-\lambda_1\exp\{-\lambda_2 y\}/\gamma_1 \text{ if } y=x \\ 1-\lambda_1\exp\{-\lambda_2 y-\lambda_3(y-x)\}/\gamma_1 \text{ if } y>x \end{cases}$$

$$F(x|y) = \begin{cases} 0 \text{ if } x<0 \\ 1-\exp\{-\lambda_1 x\} \text{ if } 0<x<y \\ 1-\lambda_2\exp\{-\lambda_1 x\}/\gamma_2 \text{ if } x=y \\ 1-\lambda_2\exp\{-\lambda_1 x-\lambda_3(x-y)\}/\gamma_2 \text{ if } x>y \end{cases}$$

As an example, $F(y|1)$ is computed below when $\lambda_1=\lambda_2=\lambda_3=1$.

$$F(y\ 1) = \begin{cases} 0 \text{ when } y<0 \\ 1-\exp\{-y\} \text{ if } 0<y<1 \\ 1-\exp\{-2y+1\}/2 \text{ if } y>1 \\ 1-\exp\{-y\}/2 \text{ if } y=1 \end{cases}$$

Note the jump at $y=1$ has size $\exp\{-1\}/2\simeq.184$.

C. Integration with Respect to the BVED.

If R is a Borel set in $(o,\infty)x(o,\infty)$ and $(X,Y)\sim BVE(\lambda_1,\lambda_2,\lambda_3)$ then

$$P[(X,Y)\varepsilon R]=\iint_R dF(x,y)=\iint_R d\overline{F}(x,y)$$

where this is meant as integration with respect to the probability measure induced by $\overline{F}(x,y;\lambda_1,\lambda_2,\lambda_3)=\exp\{-\lambda_1 x-\lambda_2 y-\lambda_3 \max(x,y)\}$. In particular if R is a rectangle, $R=\{(x,y)|a<x<b,\ c<y<d\}$, then

$$P[(X,Y)\varepsilon R]=\iint_R d\overline{F}(x,y)=\overline{F}(b,d)+\overline{F}(a,c)-\overline{F}(a,d)-\overline{F}(b,c)$$

or if $R=\{(x,y)|x>a,\ y>b\}$ then

$$P[(X,Y)\varepsilon R]=\iint_R d\overline{F}(x,y)=\overline{F}(a,b).$$

For a general set R (recall (1) and the definitions of R1, R2, R3)

$$P[(X,Y)\varepsilon R]=\iint_R d\overline{F}(x,y)= \iint_{R\cap R1} \lambda_1\gamma_2\overline{F}(x,y)dxdy +$$

$$\iint_{R\cap R2} \lambda_2\gamma_1\overline{F}(x,y)dxdy + \int_{R\cap R3} \lambda_3\overline{F}(x,x)dx.$$

That is, integration with respect to $\overline{F}$ may be represented in terms of ordinary Riemann integration. If $g(x,y)$ is a continuous function on

$(o,\infty)x(o,\infty)$ then also $\int\int g(x,y)d\overline{F}(x,y)$ over R may be expressed as the sum of three ordinary Riemann integrals:

$$\int\int_R g(x,y)d\overline{F}(x,y) = \int\int_{R\cap R1} g(x,y)\lambda_1\gamma_2\overline{F}(x,y)dxdy +$$

$$\int\int_{R\cap R2} g(x,y)\lambda_2\gamma_1\overline{F}(x,y)dxdy + \int_{R\cap R3} g(x,x)\lambda_3\overline{F}(x,x)dx. \quad (2)$$

Suppose $(X,Y)\sim BVE(\lambda_1, \lambda_2, \lambda_3)$ and $Z=\max(X,Y)$. Now $P[Z<z] = P[X<z, Y<z] = F(z,z) = \overline{F}(z,z) + \overline{F}(0,0) - \overline{F}(0,z) - \overline{F}(z,0) = 1 + \exp\{-\lambda z\} - \exp\{-\gamma_1 z\} - \exp\{-\gamma_2 z\}$ so that the probability density function of Z, call it $g(z)$, may be obtained by differentiation: $g(z)=\gamma_1\exp\{-\gamma_1 z\} + \gamma_2\exp\{-\gamma_2 z\} - \lambda\exp\{-\lambda z\}$ for $z>0$. Using $g(z)$, the values for $E(Z)$, $E(Z^2)$, $Var(Z)$ may be obtained in a straightforward manner. The values obtained are

$$E(Z) = 1/\gamma_1 + 1/\gamma_2 - 1/\lambda$$

$$E(Z^2) = 2/\gamma_1^2 + 2/\gamma_2^2 - 2/\lambda^2$$

$$Var(Z) = 1/\lambda_1^2 + 1/\lambda_2^2 - 3/\lambda^2 - 2/\gamma_1\gamma_2 + 2/\gamma_1\lambda + 2/\gamma_2\lambda$$

Without finding the p.d.f. of Z one may use (2) to obtain the values for $E(Z)$, $E(Z^2)$. Since $E(Z)=\int\int \max(x,y)d\overline{F}(x,y)$ one has

$$E(Z) = \int\int_{R1} y\lambda_1\gamma_2\overline{F}(x,y)dxdy + \int\int_{R2} x\lambda_2\gamma_1\overline{F}(x,y)dxdy + \int_{R3} x\lambda_3\overline{F}(x,x)dx.$$

Each of the three integrals can be evaluated readily though not as conveniently as when using g. The first integral on the right yields $1/\gamma_2 - \gamma_2/\lambda^2$, the second $1/\gamma_1 - \gamma_1/\lambda^2$, and the third λ_3/λ^2. Adding the terms together gives $1/\gamma_1 + 1/\gamma_2 - 1/\lambda$ (since $\gamma_1+\gamma_2-\lambda_3=\lambda$), which is the same result as above. The computation for $E(Z^2)$ is similar.

The distribution of X+Y where $(X,Y)\sim BVE(\lambda_1, \lambda_2, \lambda_3)$ is found as another example of this manner of integration. Let $Z=X+Y$ and $G(z)=P[Z<z]=P[X+Y<z]=\int\int_R d\overline{F}(x,y)$ where $R=\{(x,y)|x+y<z \text{ and } x>0, y>0\}$.

Using (2)

$$G(z) = \iint_{R\cap R1} \lambda_1\gamma_2\overline{F}(x,y)dxdy + \iint_{R\cap R2} \lambda_2\gamma_1\overline{F}(x,y)dxdy + \int_{R\cap R3} \lambda_3\overline{F}(x,x)dx.$$

Or specifically

$$G(z) = \int_0^{z/2} \lambda_1\exp\{-\lambda_1 x\} \int_x^{z-x} \gamma_2\exp\{-\gamma_2 y\}dydx + \int_0^{z/2} \lambda_2\exp\{-\lambda_2 y\} \int_y^{z-y} \gamma_1\exp\{-\gamma_1 x\}dxdy$$
$$+ \int_0^{z/2} \lambda_3\exp\{-\lambda x\}dx.$$

These integrations may be carried out with little trouble. The distributions of X-Y and X/Y may be obtained in this manner also.

In Appendix C certain expected values are required. For example, the expected value of X over R1 is required. This will be denoted by $\underset{R1}{E}(X)$ and is computed by $\iint_{R1} x\lambda_1\gamma_2\overline{F}(x,y)dxdy$. It is found that $\underset{R1}{E}(X) = \lambda_1/\lambda^2$ and $\underset{R2}{E}(X)=1/\gamma_1-\gamma_1/\lambda^2$ and $\underset{R3}{E}(X)=\lambda_3/\lambda^2$ so that $E(X)=1/\gamma_1=\underset{R1}{E}(X)+\underset{R2}{E}(X)+\underset{R3}{E}(X)$.

D. Moments of the Bivariate Exponential Distribution

In [1] Marshall and Olkin develop a formula for the moments about the origin of the bivariate exponential distribution. The central moments through the fourth order are tabulated here for reference and use in a later portion of this thesis. Let $E[(X-a)^i(Y-b)^j]$ be denoted by μ_{ij} where $a=E(X)=1/\gamma_1$ and $b=E(Y)=1/\gamma_2$ and $(X,Y)\sim BVE(\lambda_1, \lambda_2, \lambda_3)$. The following relations hold:

$$\lambda=\lambda_1+\lambda_2+\lambda_3 \qquad \rho=\lambda_3/\lambda$$

$$\gamma_1=\lambda_1+\lambda_3 \qquad \gamma_2=\lambda_2+\lambda_3$$

$$\mu_{20}=1/\gamma_1^{\,2} \qquad \mu_{03}=2/\gamma_2^{\,3}$$

$$\mu_{11}=\rho/\gamma_1\gamma_2 \qquad \mu_{40}=9/\gamma_1^{\,4}$$

$\mu_{02}=1/\gamma_2{}^2$

$\mu_{30}=2/\gamma_1{}^3$

$\mu_{21}=2\rho/\gamma_1\gamma_2\lambda$

$\mu_{12}=2\rho/\lambda\gamma_1\gamma_2$

$\mu_{31}=3\rho(2/\lambda^2+1/\gamma_1{}^2)/\gamma_1\gamma_2$

$\mu_{22}=8\rho/\gamma_1\gamma_2\lambda^2 + 1/\gamma_1{}^2\gamma_2{}^2$

$\mu_{13}=3\rho(2/\lambda^2 + 1/\gamma_2{}^2)/\gamma_1\gamma_2$

$\mu_{04}=9/\gamma_2{}^4$

The marginal moments μ_{io}, μ_{oj} are easy to obtain from the marginal distributions. However, for i and j positive the computations are tedious. If $\rho=0$, that is, if $(X,Y)\sim BVE(\gamma_1, \gamma_2, 0)$ so that X and Y are independent with marginal exponentials, the moments given here will correspond to the moments for the independent case.

E. Distribution of $(\overline{X},\overline{Y})$.

In this section a comparison is made between the distribution of $(\overline{X},\overline{Y})$ based on a random sample of size n from $BVE(\lambda_1, \lambda_2, \lambda_3)$ and an appropriate bivariate normal distribution (BVN) and a bivariate gamma type distribution (BVG) given by Kibble in [9]. Since the distribution of $(\overline{X},\overline{Y})$ is hard to obtain, the comparison will be made by examining the moments of the distribution of $(\overline{X},\overline{Y})$ when $(X,Y)\sim BVE(\lambda_1,\lambda_2,\lambda_3)$ in relation to the moments of the BVN and BVG. Some Monte Carlo simulations of the distribution of $(\overline{X},\overline{Y})$ will illustrate the approximations for fixed sample size n and values of λ_1, λ_2, λ_3.

Suppose $(X,Y)\sim BVE(\lambda_1,\lambda_2,\lambda_3)$ and $\{(X_j,Y_j)\},(j=1,\ldots,n)$ is a random sample of size n. The moments for (X,Y) may be computed from $\psi(s,t) = E[\exp -sX-tY]$ where $\psi(s,t)$ is given in [1]. The central moments through the fourth order have been given in the previous section. Let $\overline{X} = \sum X_j/n$ and $\overline{Y} = \sum Y_j/n$. Then from properties of expected values $E[\exp\{-s\overline{X}-t\overline{Y}\}] = [\psi(s/n, t/n)]^n$. If this expression is expanded out the raw moments for $\bar{X}$ and $\bar{Y}$ may be computed. The raw moments and the central moments have been calculated through the third order and are tabulated in Table I.

For sufficiently large n the central limit theorem implies that $(\bar{X},\bar{Y})$ has approximately a bivariate normal distribution with mean vector $(1/\gamma_1 \; 1/\gamma_2)$ and covariance matrix

$$n^{-1}\begin{pmatrix} 1/\gamma_1 & \rho/\gamma_1\gamma_2 \\ & \\ \rho/\gamma_1\gamma_2 & 1/\gamma_2^2 \end{pmatrix}$$

Under the assumption that $(\bar{X},\bar{Y})$ has a bivariate normal distribution with mean vector and covariance matrix given above, then

$$E[\exp\{-s\bar{X}-t\bar{Y}\}] = \exp\{-s/\gamma_1-t/\gamma_2+s^2/2n\gamma_1^2+\rho st/n\gamma_1\gamma_2+t^2/2n\gamma_2^2\}$$

The first and second moments will agree exactly with those in Table I and it is known that the central third order moments are zero. The raw third moments are

$$\mu'_{30} = 3/n\gamma_1^3 + 1/\gamma_1^3$$

$$\mu'_{03} = 3/n\gamma_2^3 + 1/\gamma_2^3$$

$$\mu'_{21} = 2\rho/n\gamma_1^2\gamma_2 + 1/n\gamma_1^2\gamma_2 + 1/\gamma_1^2\gamma_2$$

$$\mu'_{12} = 2\rho/n\gamma_1\gamma_2^2 + 1/n\gamma_1\gamma_2^2 + 1/\gamma_1\gamma_2^2$$

It is apparent that for n large and ρ small there is a close agreement between the moments of $(\bar{X},\bar{Y})$ assumed from BVE with those of $(\bar{X},\bar{Y})$ assumed from BVN.

TABLE I

MOMENTS FOR $\overline{X}$ AND $\overline{Y}$ BASED ON A SAMPLE OF SIZE n

BVE$(\lambda_1,\lambda_2,\lambda_3)$

$\mu'_{10} = 1/\gamma_1$

$\mu'_{01} = 1/\gamma_2$

$\mu'_{20} = 1/n\gamma_1{}^2 + 1/\gamma_1{}^2$

$\mu'_{02} = 1/n\gamma_2{}^2 + 1/\gamma_2{}^2$

$\mu'_{11} = \rho/n\gamma_1\gamma_2 + 1/\gamma_1\gamma_2$

$\mu'_{30} = 6/n^2\gamma_1{}^3 + 6(n-1)/n^2\gamma_1{}^3 + (n-1)(n-2)/n^2\gamma_1{}^3$

$\mu'_{03} = 6/n^2\gamma_2{}^3 + 6(n-3)/n^2\gamma_2{}^3 + (n-1)(n-2)/n^2\gamma_2{}^3$

$\mu'_{21} = 2\rho(1/n\gamma_1 + 1/n\gamma_2)/\gamma_1\gamma_2 + 1/n\gamma_1{}^2\gamma_2 + 1/\gamma_1{}^2\gamma_2$

$\mu'_{12} = 2\rho(1/n\gamma_1 + 1/n\gamma_2)/\gamma_1\gamma_2 + 1/n\gamma_1\gamma_2{}^2 + 1/\gamma_1\gamma_2{}^2$

$\mu_{20} = 1/n\gamma_1{}^2$

$\mu_{02} = 1/n\gamma_2{}^2$

$\mu_{11} = \rho/n\gamma_1\gamma_2$

$\mu_{30} = 2/n^2\gamma_1{}^3$

$\mu_{03} = 2/n^2\gamma_2{}^3$

$\mu_{21} = 2\rho/n^2\lambda\gamma_1\gamma_2$

$\mu_{12} = 2\rho/n^2\lambda\gamma_1\gamma_2$

In [9] Kibble gives a bivariate gamma-type distribution given by the moment generating function $G(s,t) = E[\exp\{sU+tV\}] = [(1-s)(1-t)-st\rho]^{-p}$ where each of the random variables U and V have the same marginal gamma distribution and $\text{cor}(U,V)=\rho\geq 0$. X has a gamma distribution with parameters α and β $[X\sim G(\alpha,\beta)]$ if the probability density function of X is $f(x) = x^{\alpha-1}\exp\{-x/\beta\}/\beta^{\alpha}\Gamma(\alpha)$. The probability density function for U is given by $\phi(u) = u^{p-1}\exp\{-u\}/\Gamma(p)$; that is, $U\sim G(p,1)$. Also the joint probability density function of U and V is given by

$$\phi(u,v) = \phi(u)\phi(v)\sum_{r=o}^{\infty} \rho^r\Gamma(p)L_r(u,p)L_r(v,p)/r!\Gamma(p+r)$$

where $L_r(u,p)$ is a certain polynomial of degree r in u with coefficients dependent on p. Using just the first two terms of the series

$$\phi_2(u,v) = \phi(u)\phi(v)[1+\rho(u-p)(v-p)/2p] \qquad (L_1(u,p)=u-p).$$

Note that when $\rho=0$ the distribution is just the product of the marginal gamma distributions.

To adapt this distribution for comparison with that of $(\overline{X},\overline{Y})$ from BVE, let $U=\gamma_1 n\overline{X}$ and $V=\gamma_2 n\overline{Y}$. If $\{(X_j,Y_j)\}(j=1,\ldots,n)$ is a random sample from $\text{BVE}(\lambda_1,\lambda_2,\lambda_3)$ then the marginal distribution of $\overline{X}$ is a gamma with $\alpha=n$ and $\beta=1/n\gamma_1$ and the marginal distribution of $\overline{Y}$ is a gamma with $\alpha=n$ and $\beta=1/n\gamma_2$. Also $\text{cor}(\overline{X},\overline{Y})=\rho>0$ where $\rho=\lambda_3/\lambda$. So $U=\gamma_1 n\overline{X}\sim G(n,1)$ and $V=\gamma_2 n\overline{Y}\sim G(n,1)$ and $\text{cor}(U,V)=\rho$. From the moment generating function given above with p replaced by n the moments for U and V may be found. By adjusting then for the constants $\gamma_1 n$ and $\gamma_2 n$ the moments for $\overline{X}$ and $\overline{Y}$ may be found assuming that (U,V) follows this bivariate gamma distribution (BVG). The third order central moments for $(\overline{X},\overline{Y})$ assuming that $U=\gamma_1 n\overline{X}$ and $V=\gamma_2 n\overline{Y}$ follow the BVG with p=n and $\rho=\lambda_3/\lambda$ are given here:

$$\mu_{30} = 2/n^2\gamma_1^3 \qquad \mu_{21} = 2\rho/n^2\gamma_1^2\gamma_2$$

$$\mu_{03} = 2/n^2\gamma_2^3 \qquad \mu_{12} = 2\rho/n^2\gamma_1\gamma_2^2$$

It should be noted that not only will the moments through the second order match up with those corresponding to BVE given in Table I but also all marginal moments, that is, μ_{io}, μ_{oi}, μ'_{io}, μ'_{oi}, will match up since the marginal distributions in both cases are the same. The differences in the third order moments between BVE and BVG occur only in μ_{21} and μ_{12}.

To give some numerical illustrations of the idea of approximating the distribution of $(\overline{X},\overline{Y})$ based on a sample of size n from BVE$(\lambda_1,\lambda_2,\lambda_3)$, several samples of size n were simulated from BVE$(\lambda_1,\lambda_2,\lambda_3)$ and percentage points of $\overline{X}$ and $\overline{Y}$ were obtained. Interest was concentrated on the normalized variables $A=(\overline{X}-E\overline{X})/\sigma_{\overline{X}}$ and $B=(\overline{Y}-E\overline{Y})/\sigma_{\overline{Y}}$. The simulated probabilities $G_1(h)=P[|A|<h, |B|<h]$ were found for h=1.0(.2)3.0. That is, $G_1(h)$ is the probability that the normalized variables lie in a square centered at (0,0) with side 2h. These simulated values are compared with the bivariate normal probabilities [11] and the bivariate gamma probabilities. Tables of the incomplete gamma function [12] were used to evaluate the bivariate gamma probabilities and only the first two terms $\phi_2(u,v)$ for $\phi(u,v)$ were used. Table II below gives some examples of the numerical comparisons. G_2 denotes the corresponding values for BVN and G_3 denotes the corresponding values for BVG.

TABLE II

PERCENTAGE POINTS OF THE DISTRIBUTION OF $(\overline{X},\overline{Y})$

$\rho = .30$

	$G_1(h)$				$G_2(h)$	$G_3(h)$			
h	n=10	n=20	n=40	n=100		n=10	n=20	n=40	n=100
1.0	.500	.482	.476	.479	.477	.478	.471	.469	.467
1.2	.632	.605	.599	.596	.603	.611	.602	.597	.595
1.4	.741	.714	.711	.734	.711	.727	.714	.708	.706
1.6	.824	.793	.783	.811	.799	.816	.804	.797	.795
1.8	.884	.870	.860	.871	.865	.880	.870	.865	.863
2.0	.923	.914	.910	.918	.913	.920	.916	.913	.912
2.2	.947	.946	.936	.953	.946	.946	.946	.945	.946
2.4	.962	.965	.961	.977	.968	.962	.966	.966	.967
2.6	.973	.977	.973	.983	.982	.974	.976	.979	.981
2.8	.982	.985	.986	.997	.990	.981	.985	.987	.988
3.0	.987	.988	.989	.998	.994	.986	.990	.992	.994

TABLE II Continued

$\rho = .10$

h	$G_1(h)$ n=10	$G_1(h)$ n=20	$G_1(h)$ n=40	$G_1(h)$ n=100	$G_2(h)$	$G_3(h)$ n=10	$G_3(h)$ n=20	$G_3(h)$ n=40	$G_3(h)$ n=100
1.0	.480	.470	.456	.462	.467	.477	.471	.469	.467
1.2	.608	.603	.583	.590	.594	.610	.601	.597	.594
1.4	.720	.716	.697	.698	.704	.726	.713	.708	.705
1.6	.809	.807	.789	.789	.793	.814	.803	.797	.795
1.8	.874	.870	.865	.860	.862	.878	.869	.865	.863
2.0	.917	.916	.912	.917	.911	.919	.915	.913	.912
2.2	.944	.946	.943	.950	.945	.945	.945	.945	.946
2.4	.963	.964	.967	.970	.967	.961	.965	.966	.967
2.6	.973	.977	.979	.984	.981	.973	.976	.979	.981
2.8	.983	.985	.988	.990	.990	.981	.985	.987	.988
3.0	.989	.990	.992	.995	.994	.986	.990	.992	.994

IV. MAXIMUM LIKELIHOOD ESTIMATION

Suppose a random sample of size n is obtained from BVE$(\lambda_1,\lambda_2,\lambda_3)$. The equations for determining the maximum likelihood estimates $\hat{\lambda}_1$, $\hat{\lambda}_2,\hat{\lambda}_3$ of the parameters are found to be

$$\begin{aligned} n/\hat{\lambda}_1 + n_2/(\hat{\lambda}_1+\hat{\lambda}_3) &= \sum X_i \\ n_1/(\hat{\lambda}_2+\hat{\lambda}_3) + n_2/\hat{\lambda}_2 &= \sum Y_i \\ n_1/(\hat{\lambda}_2+\hat{\lambda}_3) + n_2/(\hat{\lambda}_1+\hat{\lambda}_3) + n_3/\hat{\lambda}_3 &= \sum \max(X_i,Y_i) \end{aligned} \tag{3}$$

Where n_1 counts the number of sample points where $X_i<Y_i$, n_2 counts the number of sample points where $X_i>Y_i$, and n_3 counts the number of sample points where $X_i=Y_i$. Note that $n_1 + n_2 + n_3 = n$.

The derivation of equations (3) is shown in Appendix A. It is shown in Appendix B that the BVED satisfies the usual regularity conditions, so that it follows that the sequence of solutions $(\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3)_n$ of (3) converges almost certainly to $(\lambda_1,\lambda_2,\lambda_3)$ as $n\to\infty$. Also it follows that $(\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3)$ is asymptotically distributed for large n according to the trivariate normal distribution with mean $(\lambda_1,\lambda_2,\lambda_3)$ and covariance matrix Q, where Q is the inverse of the information matrix. The information matrix for the distribution is given in Appendix B under the matrix name B. The information matrix based on a sample of size n will be denoted by $I_n(\lambda_1,\lambda_2,\lambda_3)$ and is just n times B. Furthermore the maximum likelihood estimator $(\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3)$ has asymptotic efficiency 1 for estimating $(\lambda_1,\lambda_2,\lambda_3)$. General theorems relating to the maximum likelihood estimates may be found in Chapter 12 of Wilks [10].

If one of n_1, n_2, n_3 is zero then the maximum likelihood equations (3) are not solvable for $\hat{\lambda}_1$, $\hat{\lambda}_2$, $\hat{\lambda}_3$. Consider the case where $n_1=0$. Let $AX = \sum X_j$, $AY = \sum Y_j$, $AZ = \sum \max(X_j, Y_j)$. With $n_1 = 0$ it is clear that $AX=AZ$. The maximum likelihood equations would be

$$n_2/(\hat{\lambda}_1+\hat{\lambda}_3) = AX$$

$$n_2/\hat{\lambda}_2 = AY$$

$$n_3/\hat{\lambda}_3 + n_2/(\hat{\lambda}_1+\hat{\lambda}_3) = AX$$

If $n_3>0$ one observes from the first and third equations that this system is inconsistent. If in addition $n_3=0$ then the first and third equations are equivalent and the solution to the above system is $\hat{\lambda}_2=n/AY$ and $\hat{\gamma}_1=n/AX$ where $\gamma_1 = \lambda_1+\lambda_3$. However estimates for λ_1, λ_2, λ_3 separately are not possible from this system.

Similarly if $n_2=0$ we have $AY=AZ$ and the following system for the maximum likelihood estimates.

$$n_1/\hat{\lambda}_1 = AX$$

$$n_1/(\hat{\lambda}_2+\hat{\lambda}_3) = AY$$

$$n_3/\hat{\lambda}_3 + n_1/(\hat{\lambda}_2+\hat{\lambda}_3) = AY$$

If $n_3>0$ this system is inconsistent whereas if $n_3=0$ then the solutions are $\hat{\lambda}_1 = n/AX$ and $\hat{\gamma}_2 = n/AY$ where $\gamma_2 = \lambda_2+\lambda_3$, but separate estimates are not available.

Now consider the case where $n_3=0$. The maximum likelihood equations would appear as below.

$$n_1/\hat{\lambda}_1 + n_2/(\hat{\lambda}_1+\hat{\lambda}_3) = AX \qquad \text{(a)}$$

$$n_1/(\hat{\lambda}_2+\hat{\lambda}_3) + n_2/\hat{\lambda}_2 = AY \qquad \text{(b)}$$

$$n_1/(\hat{\lambda}_2+\hat{\lambda}_3) + n_2/(\hat{\lambda}_1+\hat{\lambda}_3) = AZ \qquad \text{(c)}$$

It is not as obvious that the above system is not solvable for $\hat{\lambda}_1$, $\hat{\lambda}_2$, $\hat{\lambda}_3$. However, consider $n_1n_2>0$ and subtract equations (a) and (b) respectively from equation (c). Note that AZ>AX and AZ>AY in this case.

$$n_1[1/(\hat{\lambda}_2+\hat{\lambda}_3) - 1/\hat{\lambda}_1] = AZ-AX>0 \qquad \text{(d)}$$

$$n_2[1/(\hat{\lambda}_1+\hat{\lambda}_3) - 1/\hat{\lambda}_2] = AZ-AY>0 \qquad \text{(e)}$$

Equation (d) implies $\hat{\lambda}_1>\hat{\lambda}_2+\hat{\lambda}_3$ and equation (e) implies $\hat{\lambda}_2>\hat{\lambda}_1+\hat{\lambda}_3$. Together these inequalities imply $\hat{\lambda}_1>(\hat{\lambda}_1+\hat{\lambda}_3)+\hat{\lambda}_3$ or $\hat{\lambda}_3<0$ which is impossible.

If both n_1 and n_3 are zero then AX=AZ and equations (a) and (c) are equivalent. The solution to the system is $\hat{\lambda}_2=n/AY$ and $\hat{\lambda}_1=n/AX$. If both n_2 and n_3 are zero then AY=AZ and equations (b) and (c) are equivalent. The solution to the system in this case is $\hat{\lambda}_1=n/AX$ and $\hat{\gamma}_2=n/AY$. So estimates for λ_1, λ_2, λ_3 separately are not available from the system when $n_3=0$ and one of n_1 or n_2 is zero.

For sample size n and λ_1, λ_2, λ_3 each positive it is clear that the probability of obtaining $n_1n_2n_3=0$ will go to zero as n increases since $P[n_i=0] = (1-\lambda_i/\lambda)^n$ and $\lambda_i<\lambda$. However for small sample size there is the possibility of obtaining an $n_i=0$. In this situation the following procedure is suggested.

If $n_1=0$ set $\hat{\lambda}_1=0$, $\hat{\lambda}_2=n_2/AY$, $\hat{\lambda}_3=n/AX$.

If $n_2=0$ set $\hat{\lambda}_1=n_1/AX$, $\hat{\lambda}_2=0$, $\hat{\lambda}_3=n/AY$.

If $n_3=0$ set $\hat{\lambda}_1=n/AX$, $\hat{\lambda}_2=n/AY$, $\hat{\lambda}_3=0$.

The motivation for the above procedure is suggested from the maximum likelihood equations resulting respectively from BVE$(0, \lambda_2, \lambda_3)$, BVE$(\lambda_1, 0, \lambda_3)$, BVE$(\lambda_1, \lambda_2, 0)$. For example, if n_3 is zero we find the estimates $\hat{\lambda}_1$, $\hat{\lambda}_2$ as if the distribution were the product of two independent exponential distributions with parameters λ_1 and λ_2.

In the general situation where $n_1n_2n_3>0$ the maximum likelihood equations (3) cannot be solved for $\hat{\lambda}_1$, $\hat{\lambda}_2$, $\hat{\lambda}_3$ in closed form, but an iterative procedure such as Newton's method should be employed. (A discussion of such a procedure is given in Appendix C.)

The maximum likelihood estimators are biased estimates of λ_1, λ_2, λ_3. However, multiplying each estimate by $n/(n+1)$ where n is the sample size corrects the bias reasonably well. (Random samples from BVE$(\lambda_1, \lambda_2, \lambda_3)$ were simulated and the maximum likelihood estimates computed for each sample.) Table III gives the average over several samples for various sample sizes and values of λ_1, λ_2, λ_3.

The information matrix $I_n(\lambda_1, \lambda_2, \lambda_3)$ for the distribution based on a sample of size n is nB where B is given in Appendix B. Let Q_n denote the inverse of the matrix nB. As a measure of efficiency of the estimate $(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3)$ the trace of the matrix Q_n will be used. That is, the sum of the variances or mean square errors of estimates for λ_1, λ_2, λ_3 based on a sample of size n will be compared with the trace of the Q_n matrix.

TABLE III

BIAS IN MAXIMUM LIKELIHOOD ESTIMATES

Parameters			Sample Size	Simulated Values					
λ_1	λ_2	λ_3	n	$\frac{n}{n+1} E(\hat{\lambda}_1)$	$\frac{n}{n+1} E(\hat{\lambda}_2)$	$\frac{n}{n+1} E(\hat{\lambda}_3)$	$E(\hat{\lambda}_1)$	$E(\hat{\lambda}_2)$	$E(\hat{\lambda}_3)$
.818	.818	.183	10	.837	.833	.186	.921	.916	.205
			20	.821	.820	.181	.863	.861	.190
			40	.828	.821	.183	.847	.842	.187
			100	.829	.820	.179	.837	.828	.181
			200	.820	.819	.180	.824	.823	.181
.539	.539	.461	10	.551	.547	.464	.606	.602	.511
			20	.542	.540	.460	.569	.567	.483
			40	.542	.541	.463	.555	.555	.475
			100	.545	.540	.459	.550	.545	.464
.905	.905	.095	10	.923	.917	.094	1.015	1.009	.103
			20	.904	.907	.096	.949	.952	.101
			40	.911	.910	.098	.934	.933	.100
			100	.906	.907	.095	.915	.916	.096
.364	.864	.136	10	.368	.871	.138	.405	.958	.152
			20	.369	.870	.138	.388	.914	.145
			40	.368	.864	.136	.377	.885	.140
			100	.362	.863	.137	.366	.871	.138

It is found that the trace of the matrix $Q=nQ_n$ is given by

$$Tr(Q) = \frac{(b+d)(a+d+e) - b^2 + (a+c)(a+b+e) - a^2 + (a+c)(b+d)}{(a+c)[(b+d)(a+b+e) - b^2] - a^2(b+d)} \qquad (4)$$

where $a=\lambda_2/\lambda(\lambda_1+\lambda_3)^2$, $b=\lambda_1/\lambda(\lambda_2+\lambda_3)^2$, $c=1/\lambda_1\lambda$, $d=1/\lambda_2\lambda$, $e=1/\lambda_3\lambda$ and $\lambda=\lambda_1+\lambda_2+\lambda_3$. The efficiency of the maximum likelihood estimators is compared with two other methods in Chapter V. Selected parameter sets were chosen and samples were simulated (See Table VII).

Given a restriction on the parameters of the distribution the maximum likelihood equations may be derived for that restriction and the inverse of the information matrix for the estimates may be computed. One important example is the restriction that $\lambda_1=\lambda_2$ to which reference will be made later. That is, suppose $(X,Y)\sim BVE(\lambda_1, \lambda_1, \lambda_3)$ and a random sample of size n is obtained. The maximum likelihood equations are given below as well as the solutions in closed form.

$$(n_1+n_2)/(\hat\lambda_1+\hat\lambda_3) + (n_1+n_2)/\hat\lambda_1 = AX+AY$$

$$(n_1+n_2)/(\hat\lambda_1+\hat\lambda_3) + n_3/\hat\lambda_3 = AZ$$

$$\hat\lambda_1 = [-a_2 - (a_2 - 4a_1a_3)^{1/2}]/2a_1$$

$$\hat\lambda_3 = \hat\lambda_1[(AX+AY)\hat\lambda_1 - 2(n_1+n_2)]/[(n_1+n_2) - (AX+AY)\hat\lambda_1]$$

Here $a_1 = (AX+AY)(AX+AY-AZ)$, $a_2 = -2(n_1+n_2)(AX+AY-AZ) - n(AX+AY)$ and $a_3 = (n_1+n_2)(n+n_1+n_2)$.

The inverse of the information matrix $I_n(\lambda_1, \lambda_3)$ is given by

$$Q_n' = (n\Delta)^{-1} \begin{pmatrix} 2b+e & -2b \\ -2b & 2b+2c \end{pmatrix}$$

where $\Delta = 2b(2c+e) + 2ce$ and b, c, and e have the same definitions as given on page 28 of this chapter.

The table below indicates the efficiency for various sample sizes and for selected values of the parameters. $Tr(Q_n')$ denotes the trace of the matrix Q_n'.

TABLE IV

EFFICIENCY OF MAXIMUM LIKELIHOOD ESTIMATES WHEN $\lambda_1 = \lambda_2$

Parameters $\lambda_1 = \lambda_2$	λ_3	Sample Size n	$Tr(Q_n')$	Simulated Values MSE(MLE)	Efficiency
.818	.182	10	.0771	.0797	.97
		20	.0386	.0380	.98
		40	.0193	.0195	.99
		100	.0077	.0077	1.00
		200	.0039	.0039	1.00
.539	.461	10	.0888	.0925	.96
		20	.0444	.0458	.97
		40	.0222	.0227	.98
		100	.0089	.0089	1.00
.905	.095	10	.0673	.0688	.98
		20	.0337	.0339	.99
		40	.0168	.0171	.98
		100	.0084	.0084	1.00

V. TWO OTHER METHODS FOR PARAMETER ESTIMATION

A. Introduction

In this chapter a method to estimate the parameters will be given which is motivated by the classical method of moments and by the particular nature of the bivariate exponential distribution. It will be seen that this method (which will be referred to as PROP) is a good competitor with the method of maximum likelihood estimation (MLE) given in Chapter IV. The PROP method has the advantage of ease of computation compared to the MLE method and yet has efficiency close to MLE.

The second method of estimation to be discussed here has been given by Arnold in [2]. This method will be referred to as ACE for Arnold's consistent estimates. The estimates given by ACE have the advantage of being unbiased and share the feature of ease of computation with PROP but do not compete well with MLE or PROP in terms of efficiency. The reason for this will be indicated. Arnold gave his estimates which are unbiased and consistent for the multivariate exponential distribution whereas here only estimation for the bivariate exponential distribution is considered. Finally, comparisons of these methods will be given in terms of efficiency.

B. PROP Method

In applying the idea of the classical method of moments to the bivariate exponential distribution we seek three statistics from a random sample of size n drawn from BVE(λ_1, λ_2, λ_3) whose expected values involve the three parameters of the distribution. Equating the statistics with their expected values and solving for the parameters in terms of the statistics we obtain estimates of the parameters. The choice of which three statistics to use does not seem to be unique. The natural choice would seem to be $\sum X_j$, $\sum Y_j$, and $\sum X_jY_j$. However, the following three seem to do better than others that have been tried thus far. They are $\sum X_j$, $\sum Y_j$, and n_3; that is, the sum of the X observations, the sum of the Y observations and the number of sample points for which $X_j=Y_j$. Since X and Y both have marginal exponential distributions with parameters $\lambda_1+\lambda_3$ and $\lambda_2+\lambda_3$, respectively, $E(\sum X_j) = n/(\lambda_1+\lambda_3)$ and $E(\sum Y_j) = n/(\lambda_2+\lambda_3)$. The number, n_3, of sample points where $X_j=Y_j$ is a random variable which has a binomial distribution with parameters n and λ_3/λ where $\lambda=\lambda_1+\lambda_2+\lambda_3$. Also the correlation between X and Y in the BVED is precisely λ_3/λ as shown in [1]. Now $E(n_3) = n\rho = n\lambda_3/\lambda$. In terms of moments the marginal means and the correlation are used. The equations for the PROP estimates are summarized below.

$$\begin{aligned} \sum X_j &= n/(\hat{\lambda}_1+\hat{\lambda}_3) \\ \sum Y_j &= n/(\hat{\lambda}_2+\hat{\lambda}_3) \\ n_3 &= n\hat{\lambda}_3/(\hat{\lambda}_1+\hat{\lambda}_2+\hat{\lambda}_3) \end{aligned}$$

These equations may be readily solved for $\hat{\lambda}_1$, $\hat{\lambda}_2$, $\hat{\lambda}_3$ and the solutions are given below.

$$\begin{aligned} \hat{\lambda}_1 &= (n/AX-n_3/AY)/(1+n_3/n) \\ \hat{\lambda}_2 &= (n/AY-n_3/AX)/(1+n_3/n) \\ \hat{\lambda}_3 &= n_3(1/AX+1/AY)/(1+n_3/n) \end{aligned} \tag{5}$$

where $AX=\sum X_j$ and $AY=\sum Y_j$. Note that when $n_3=0$, PROP yields $\hat{\lambda}_3=0$, $\hat{\lambda}_1=n/AX$, $\hat{\lambda}_2=n/AY$.

The usual sample correlation coefficient

$$r = \sum(X_j-\overline{X})(Y_j-\overline{Y})/[\sum(X_j-\overline{X})^2\sum(Y_j-\overline{Y})^2]^{1/2}$$

used in normal theory is not a good estimator of ρ in the bivariate exponential distribution. Instead a natural estimate of ρ is n_3/n, the proportion of the sample points such that $X_j=Y_j$, due to the particular feature of the BVED that $P[X=Y] = \rho$. The discussion below indicates that n_3/n is a much more efficient estimate of ρ than r for the BVED.

An approximate result for the variance of r has been given on page 359 in Cramér's book [13] in terms of the central moments of the parent distribution which has error of order $n^{-3/2}$ where n is the sample size on which r is based. This result is repeated here where $D^2(r)$ denotes the approximate variance of r:

$$D^2(r) = \rho^2(\mu_{40}/\mu_{20}^2 + \mu_{04}/\mu_{02}^2 + 2\mu_{22}/\mu_{20}\mu_{02} + 4\mu_{22}/\mu_{11}^2$$
$$- 4\mu_{31}/\mu_{11}\mu_{20} - 4\mu_{13}/\mu_{11}\mu_{02})/4n$$

Using the moments given in Chapter III for the BVED then

$$D^2(r) = \{(1-\rho^2) + 2[2\gamma_1\gamma_2(2+\rho^2)\rho - 3\rho^2(\gamma_1{}^2+\gamma_2{}^2)]/\lambda^2\}/n.$$

In the special case when $\gamma_1=\gamma_2=\gamma$ the result simplifies to $D^2(r) = (1-\rho^2)(1+2\rho+\rho^2-\rho^3)/n$.

Since n_3 has a binomial distribution with parameters n and ρ it follows that $\mathrm{Var}(n_3/n) = \rho(1-\rho)/n$. By examining the ratio of $\mathrm{Var}(n_3/n)$ to $D^2(r)$ an asymptotic efficiency of r relative to n_3/n may be obtained. In the case when $\gamma_1=\gamma_2=\gamma$ this ratio is $\rho/(1+3\rho+3\rho^2-\rho^4)$. For ρ close to zero the efficiency of r relative to n_3/n is approaching zero. The maximum of the ratio is 1/6 when $\rho=1$. By evaluating the above ratio one may compare how the efficiency increases with ρ. The agreement between simulated runs and the variances given here are close. For example, when $\rho=.1$ the asymptotic relative efficiency (A.R.E.) is 7.5% and the simulated efficiency is about 8%. Table V gives some comparisons.

Let the random variable U take on the value 0 or 1 according to whether $X \neq Y$ or $X=Y$. Then the statistic n_3 can be represented as the sum $\sum U_j$. Let Z denote the vector (X,Y,U). Then E(Z) = $(1/(\lambda_1+\lambda_3) \quad 1/(\lambda_2+\lambda_3) \quad \lambda_3/\lambda)$ and

$$\mathrm{Cov}(Z) = \begin{pmatrix} 1/(\lambda_1+\lambda_3)^2 & \lambda_3/\lambda(\lambda_1+\lambda_3)(\lambda_2+\lambda_3) & -\lambda_2\lambda_3/\lambda^2(\lambda_1+\lambda_3) \\ \lambda_3/\lambda(\lambda_1+\lambda_3)(\lambda_2+\lambda_3) & 1/(\lambda_2+\lambda_3)^2 & -\lambda_1\lambda_3/\lambda^2(\lambda_2+\lambda_3) \\ -\lambda_2\lambda_3/\lambda^2(\lambda_1+\lambda_3) & -\lambda_1\lambda_3/\lambda^2(\lambda_2+\lambda_3) & \lambda_3(\lambda_1+\lambda_2)/\lambda^2 \end{pmatrix}$$

The upper two by two matrix in Cov(Z) is just the covariance matrix of X and Y. The term in the bottom diagonal position is the variance of the point binomial random variable U. To indicate how the other terms are obtained consider the row 1, column 3 term which is the covariance of X and U.

$$E(UX) = E(U\,E_{X|U}(X)) = E_U(U\,H(U)) = 0H(0)P[U=0] + 1H(1)P[U=1]$$

Now $P[U=1] = \lambda_3/\lambda$ and H(1) denotes the expected value of X given that U=1 or X=Y.

$$H(1) = \int_0^\infty \lambda x \exp\{-\lambda x\}\, dx = 1/\lambda$$

TABLE V

COMPARISON OF n_3/n WITH r

γ_1	γ_2	ρ	n	$\rho(1-\rho)/n$	$D^2(r)$	A.R.E.	Simulated Values: $E(n_3/n)$	E(r)	$MSE(n_3/n)$	MSE(r)	Eff.
1.0	1.0	.05	10	.00480	.1100	.0432	.0489	.0472	.00474	.1154	.041
			20	.00238	.0550	.0432	.0505	.0503	.00230	.0547	.042
			40	.00119	.0275	.0432	.0510	.0465	.00117	.0267	.044
			100	.00048	.0110	.0432	.0500	.0496	.00043	.0104	.041
1.0	1.0	.10	10	.00900	.1200	.0752	.0996	.1061	.00914	.1193	.077
			20	.00450	.0598	.0752	.0998	.1026	.00453	.0589	.077
			40	.00225	.0299	.0752	.0996	.1051	.00224	.0299	.075
			100	.00090	.0120	.0752	.0982	.0941	.00090	.0112	.080
1.0	1.0	.30	10	.02100	.1510	.139	.296	.320	.02071	.1214	.171
			20	.01050	.0757	.139	.298	.318	.01032	.0637	.162
			40	.00525	.0378	.139	.299	.308	.00539	.0363	.148
			100	.00210	.0151	.139	.298	.304	.00207	.0128	.162
.5	1.0	.10	10	.09000	.1166	.0772	.1002	.1109	.00899	.1151	.078
			20	.00450	.0584	.0772	.0998	.1047	.00455	.0569	.080
			40	.00225	.0292	.0772	.0995	.1112	.00221	.0289	.076
			100	.00090	.0117	.0772	.1005	.1024	.00097	.0125	.078

Hence $E(UX) = \lambda_3/\lambda^2$ and thus

$$\mathrm{Cov}(U,X) = E(UX) - E(U)E(X) = \lambda_3/\lambda^2 - \lambda_3/\lambda(\lambda_1+\lambda_3) =$$

$$\lambda_3(\lambda_1+\lambda_3-\lambda)/\lambda^2(\lambda_1+\lambda_3) = -\lambda_2\lambda_3/\lambda^2(\lambda_1+\lambda_3).$$

$\mathrm{Cov}(U,Y)$ is obtained similarly.

If $\{(X_j,Y_j)\}$ is a random sample of size n from $\mathrm{BVE}(\lambda_1,\lambda_2,\lambda_3)$ then by the central limit theorem $\sum[Z_j-E(Z)]/n^{1/2}$ is approximately distributed as a trivariate normal with mean (0,0,0) and covariance matrix Cov(z). Or $\sum Z_j/n$ is approximately trivariate normal with mean E(Z) and covariance matrix Cov(Z)/n.

For notational convenience let $a=\sum X_j/n$, $b=\sum Y_j/n$, $c=n_3/n=\sum U_j/n$. Then the equations (5) may be written in the form

$$\hat{\lambda}_1 = (1/a-c/b)/(1+c)$$

$$\hat{\lambda}_2 = (1/b-c/a)/(1+c)$$

$$\hat{\lambda}_3 = c(1/a+1/b)/(1+c)$$

Also $\sum Z_j/n = (a,b,c)$. Let $(a_o,b_o,c_o) = E(Z)$; that is, $a_o = 1/(\lambda_1+\lambda_3)$, $b_o = 1/(\lambda_2+\lambda_3)$, $c_o = \lambda_3/\lambda$. If $\hat{L}$ denotes the vector of PROP estimates and L denotes the vector of parameter values then by expanding equations (5) through linear terms about (a_o,b_o,c_o), $\hat{L}-L$ is approximately $B[\sum Z_j/n-E(Z)]$ and $\hat{L}$ is approximately distributed as a trivariate normal distribution with mean L and covariance matrix $B\mathrm{Cov}(Z)B^T/n$. The matrix B is given below. Note that B is a matrix of full rank, since

$$\det(B) = (\lambda_3+\lambda)(\lambda_1+\lambda_3)^2(\lambda_2+\lambda_3)^2 \neq 0.$$

$$B = \lambda(\lambda_3+\lambda)^{-1}\begin{pmatrix} -(\lambda_1+\lambda_3)^2 & \lambda_3(\lambda_2+\lambda_3)^2/\lambda & -\lambda \\ \lambda_3(\lambda_1+\lambda_3)^2/\lambda & -(\lambda_2+\lambda_3)^2 & -\lambda \\ -\lambda_3(\lambda_1+\lambda_3)^2/\lambda & -\lambda_3(\lambda_2+\lambda_3)^2/\lambda & \lambda \end{pmatrix}$$

(Appendix D gives details of the expansion and computations.)

The matrix $BCov(Z)B^T/n$ may be used then as an approximation to the asymptotic covariance matrix for the PROP estimates. The matrix $A = (\rho+1)^2BCov(Z)B^T$ is symmetric and the entries are given below, where $\rho = \lambda_3/\lambda$.

$$a_{11} = (\lambda_1+\lambda_3)^2 + \rho[(\lambda_1+\lambda_3)^2 + \lambda_2^2] - \rho^2[\lambda_1^2+2(\lambda_1+\lambda_3)(\lambda_2+\lambda_3)]$$

$$a_{22} = (\lambda_2+\lambda_3)^2 + \rho[(\lambda_2+\lambda_3)^2 + \lambda_1^2] - \rho^2[\lambda_2^2+2(\lambda_1+\lambda_3)(\lambda_2+\lambda_3)]$$

$$a_{33} = \rho\lambda(\lambda_1+\lambda_3) + \rho^2[(\lambda_1+\lambda_3)^2+(\lambda_2+\lambda_3)^2+2(\lambda_1+\lambda_3)\lambda_2+2(\lambda_2+\lambda_3)\lambda_1] + 2\rho^3(\lambda_1+\lambda_3)(\lambda_2+\lambda_3)$$

$$a_{12} = \rho[\lambda(\lambda_1+\lambda_2) + (\lambda_1+\lambda_3)(\lambda_2+\lambda_3) - (\lambda_1+\lambda_3)^2 - (\lambda_2+\lambda_3)^2-(\lambda_1+\lambda_3)\lambda_2 -(\lambda_2+\lambda_3)\lambda_1] + \rho^2[(\lambda_1+\lambda_3)\lambda_2+(\lambda_2+\lambda_3)\lambda_1] + \rho^3(\lambda_1+\lambda_3)(\lambda_2+\lambda_3)$$

$$a_{13} = \rho[(\lambda_1+\lambda_3)^2+(\lambda_1+\lambda_3)\lambda_2-\lambda(\lambda_1+\lambda_2)] + \rho^2[(\lambda_1+\lambda_3)(\lambda_2+\lambda_3)- (\lambda_2+\lambda_3)^2-(\lambda_1+\lambda_3)\lambda_2-2\lambda_1(\lambda_2+\lambda_3)] - \rho^3(\lambda_1+\lambda_3)(\lambda_2+\lambda_3)$$

$$a_{23} = \rho[(\lambda_2+\lambda_3)^2 + (\lambda_2+\lambda_3)\lambda_1 - \lambda(\lambda_1+\lambda_2)] - \rho^3(\lambda_1+\lambda_3)(\lambda_2+\lambda_3) + \rho^2[(\lambda_1+\lambda_3)(\lambda_2+\lambda_3) - (\lambda_1+\lambda_3)^2- (\lambda_2+\lambda_3)\lambda_1 - 2\lambda_2(\lambda_1+\lambda_3)]$$

For various parameter sets the trace of the inverse of the information matrix $I_n(\lambda_1,\lambda_2,\lambda_3)$ times the sample size n (trace of Q in

Chapter IV), the trace of the matrix $BCov(Z)B^T$, which serves as an asymptotic lower bound for the sum of the mean square errors of the PROP estimates of the parameters, and the results of simulations may be compared. Also by comparing the terms on the main diagonal of $BCov(Z)B^T$ with those on the main diagonal of Q, one may compare the efficiencies of the individual estimates by the PROP method. This will be done in the last section of this chapter.

If the equations (5) for the PROP estimates are expanded through quadratic terms about (a_o, b_o, c_o) one can see that the estimates will be biased. However, multiplying the $\hat{\lambda}_i$ given by equations (5) by the factor n/(n+1) does a reasonable job of correcting for this bias. Appendix D gives the details.

Suppose it is known that the marginal parameters γ_1 and γ_2 are equal (equivalently $\lambda_1=\lambda_2$) then the PROP estimation procedure for λ_1, λ_2, λ_3 is modified to take into account this information. The equations for the PROP estimates used in this case are

$$2n/(\hat{\lambda}_1+\hat{\lambda}_3) = AX+AY$$

$$n\hat{\lambda}_3/(2\hat{\lambda}_1+\hat{\lambda}_3) = n_3$$

Since $E(AX) = E(AY) = n/(\lambda_1+\lambda_3) = n/\gamma_1$, the first equation above comes from equating E(AX+AY) to AX+AY with the parameters replaced by estimates. The second equation is the same as before arising from $E(n_3)$ except using $\lambda_1=\lambda_2$ in this case. The solutions for the estimates are

$$\hat{\lambda}_1 = 2n(n_1+n_3)/(AX+AY)(n+n_3)$$

$$\hat{\lambda}_3 = 4n\ n_3/(AX+AY)(n+n_3)$$

C. ACE

Suppose $\{(X_j,Y_j)\}$ is a random sample of size n from $BVE(\lambda_1,\lambda_2,\lambda_3)$. In [2] the statistics n_1,n_2, n_3 and $\sum\min(X_j,Y_j)=AU$ are used to estimate the parameters λ_1, λ_2, λ_3 of the BVED in the following way.

$$\hat{\lambda}_1 = n_1(n-1)/nAU$$

$$\hat{\lambda}_2 = n_2(n-1)/nAU \tag{6}$$

$$\hat{\lambda}_3 = n_3(n-1)/nAU$$

It was shown that the estimates given by equations (6) are unbiased and consistent and that the sum of the mean square errors $\sum MSE(\hat{\lambda}_i)$ $(i=1,2,3)$ is given by $[(n-1)\lambda^2+\sum\lambda_i^2]/n(n-2)$. This last expression will be used to compare the efficiency of ACE with MLE and PROP in Section D. Also in [2] it is pointed out that AU is uncorrelated with each of n_1, n_2, n_3 which will be used later in Chapters VI and VII.

It is shown here that indeed if one has available only those statistics n_1, n_2, n_3 and AU, then the estimates given by equations (6) are the minimum variance unbiased estimates for λ_1, λ_2, λ_3.

Suppose $(X,Y)\sim BVE(\lambda_1, \lambda_2, \lambda_3)$ and let $U_1= 1$ if $X<Y$ and 0 otherwise, let $U_2= 1$ if $X>Y$ and 0 otherwise, let $U=\min(X,Y)$. Now U has an exponential distribution with parameter λ; that is, U has density $\lambda\exp\{-\lambda u\}$ with respect to the usual Lebesgue measure on $(0, \infty)$. The random vector $V=(U_1, U_2)$ takes on values (1,0), (0,1), (0,0) with probabilities λ_1/λ, λ_2/λ, λ_3/λ, respectively. If $p_1=\lambda_1/\lambda$ and $p_2=\lambda_2/\lambda$ then V has density

$$p_1^{u_1}\, p_2^{u_2}\, (1-p_1-p_2)^{1-u_1-u_2}$$

with respect to the counting measure on the space $\{(0,0), (0,1), (1,0)\}$.

Let $M=(0,\infty)$ and M^* denote the σ-algebra of Borel sets which are subsets of M and μ denote the usual Lebesgue measure on M^*. Then (M,M^*,μ) is a σ-finite measure space. Let $N=\{(0,0), (0,1), (1,0)\}$ and N^* denote the power set of N and ν denote the counting measure on N^*. Then (N,N^*,ν) is a finite measure space.

The joint density of (U,V) will be found with respect to the product measure on MxN. Let $S=MxN$ and S^* denote the σ-algebra on S generated by the class of product sets, the first from M^* and the second from N^*. Define the measure σ on S^* by $\sigma(A_M \times A_N) = \mu(A_M) \times \nu(A_N)$ where $A_M \varepsilon M^*$ and $A_N \varepsilon N^*$. It is known that σ defined on rectangle sets in the above way uniquely determines the product measure $\sigma = \mu \times \nu$ on S^*. [14]

Let A be a set in S^*. $\sigma(A) = \mu \times \nu\ (A) = \int\mu(A_n)d\nu = \int\nu(A_m)d\mu$ where A_n is the section of A at the point $n \varepsilon N$ and A_m is the section of A at the point $m \varepsilon M$. If $\sigma(A)$ were zero then $\mu(A_n)$ must be zero for $n \varepsilon N$, implying each of the sets A_n in M^* have measure zero or $A_m=\phi$ for each $m \varepsilon M$. In either case it is clear that $\sigma(A)=0$ implies $P(A)=0$ where P denotes the probability measure associated with (U,V). Hence the probability measure for (U,V) on S^* is absolutely continuous with respect to the product measure σ. By the Radon-Nikodym theorem then it makes sense to talk about the joint density of (U,V) with respect to σ. The obvious choice for this density is

$$f(u,u_1,u_2) = \lambda\exp\{-\lambda u\}p_1^{u_1}\ p_2^{u_2}\ (1-p_1-p_2)^{1-u_1-u_2}$$

in view of the fact that U and V are independent.

Since the joint density given by $f(u,u_1,u_2)$ above has the form of a density in the exponential class

$$f(u,u_1,u_2) = \lambda(1-p_1-p_2)\exp\{-\lambda u + u_1\log[p_1/(1-p_1-p_2)] + u_2\log[p_2/(1-p_1-p_2)]\}$$

one may inspect for the sufficient statistics. Let $\{(U_j, U_{1j}, U_{2j})\}$ be a random sample of size n and $AU = \sum U_j$, $n_1 = \sum U_{1j}$, $n_2 = \sum U_{2j}$. Then by the factorization theorem [15] AU, n_1, n_2 are sufficient for λ, p_1, p_2. Equivalently AU, n_1, n_2 are sufficient for λ_1, λ_2, λ_3 since there is a 1-1 transformation from (λ, p_1, p_2) to $(\lambda_1, \lambda_2, \lambda_3)$.

Consider the maximum likelihood estimates for λ_1, λ_2, λ_3 in this restricted sampling setting. If $L = \Pi f(U_j, U_{1j}, U_{2j})$ then $\log(L) = n\log(\lambda)-\lambda AU+n_1\log(p_1)+n_2\log(p_2)+n_3\log(1-p_1-p_2)$ where $n_3 = n-n_1-n_2$. Taking partials with respect to λ_1, λ_2, λ_3 and equating to zero the estimates are $\hat{\lambda}_1 = n_1/AU$, $\hat{\lambda}_2 = n_2/AU$, $\hat{\lambda}_3 = n_3/AU$ (also $\hat{\lambda} = n/AU$). The maximum likelihood estimates are functions of the sufficient statistics. If they are unbiased then the unique minimum variance unbiased estimates are obtained. Since $E(1/AU) = \lambda/(n-1)$ and n_1, n_2 are independent of AU we obtain as the minimum variance unbiased estimates for λ_1, λ_2, λ_3 the same estimates as given by Arnold in [2], namely those given in equations (6). In an abstract [8] Maik makes reference to this also.

In just considering the statistics n_1, n_2, n_3, $\sum\min(X_j,Y_j)$ one does not use all the information from the sample. In a life test situation both X_j and Y_j may not be observable if $X_j \neq Y_j$. In this case the estimates given by (6) are available to use. However, observing all X_j and Y_j so that $\sum X_j$, $\sum Y_j$ and $\sum \max(X_j, Y_j)$ are available indicates why PROP and MLE are better in general to use than ACE. In computing the information matrix $I_n(\lambda_1, \lambda_2, \lambda_3)$ obtained by Arnold in [2] and also in Appendix B here, the availability of both responses X_j and Y_j for each sample point is tacitly assumed.

Arnold does point out in [2] that the efficiency of his estimates range from .5 to 1.0 depending on the parameter values. If the parameter values are $\lambda_1 = \lambda_2 = a$, $\lambda_3 = 1$ and a becomes large, the efficiency of the ACE estimates approaches 50%. But this may be one of the more interesting situations in which the BVED should be applied; that is, when the marginal parameters for X and Y are equal and there is a small positive correlation. Examples of this situation will be considered in later chapters.

Suppose it is known that the marginal parameters γ_1 and γ_2 are equal (equivalently $\lambda_1 = \lambda_2$) then the ACE estimation procedure for $\lambda_1, \lambda_2, \lambda_3$ is modified to take into account this information. The equations for the ACE estimates used in this case are

$$\hat{\lambda}_1 = (n_1+n_2)(n-1)/2nAU$$

$$\hat{\lambda}_3 = n_3(n-1)/nAU.$$

Note that the estimate for λ_3 remains the same as when no information about $\lambda_1 = \lambda_2$ is available and the estimate for λ_1 has been modified by using $(n_1+n_2)/2$ in place of n_1 in the formula for $\hat{\lambda}_1$ given previously in (6), or using $(n_1+n_2)/2$ in place of n_2 in the formula for $\hat{\lambda}_2$ given in (6) since here $\hat{\lambda}_1 = \hat{\lambda}_2$.

D. Comparison

Both MLE and PROP yield biased estimates of the parameters whereas ACE yields unbiased estimates. All three methods yield consistent estimates. The efficiency of the estimation methods is given in terms of the ratio of the trace of the inverse of the information matrix to the sum of the mean square errors of $\hat{\lambda}_1$, $\hat{\lambda}_2$, and $\hat{\lambda}_3$. For the trace of the inverse of the information matrix the quantity used is Tr(Q) given by equation (4) in Chapter IV. For the sum of the mean square errors for the PROP method the trace of the matrix $BCov(Z)B^T/n$ is used:

$$\text{MSE(PROP)} = \{(\lambda_1+\lambda_3)^2 + (\lambda_2+\lambda_3)^2 + \lambda_3(\lambda_1+\lambda_2) + 2\rho^3\lambda_1\lambda_2 + \rho[(\lambda_1+\lambda_3)^2 + (\lambda_2+\lambda_3)^2 + \lambda_1^2 + \lambda_2^2]\}/n(\rho+1)^2 \qquad (7)$$

For the sum of the mean square errors for the ACE method the quantity $[(n-1)\lambda^2 + (\lambda_1^2+\lambda_2^2+\lambda_3^2)]/n(n-2)$ is given in [2]. The quantity below neglecting terms of order $O(1/n)$ is used for the comparison.

$$\text{MSE(ACE)} = \lambda^2/n \qquad (8)$$

Table VI below compares these quantities: equation (4) in Chapter IV, equations (7) and (8) above multiplied by sample size for various values of λ_1, λ_2, λ_3 and the efficiency of PROP (ratio of (4) to (7)) and the efficiency of ACE (ratio of (4) to (8)). Let $\rho = \lambda_3/\lambda$ and $\varepsilon = \lambda_1/\lambda_2$.

Table VII gives the results of simulations of the sum of the mean square errors times sample size for all three methods (MLE, PROP, ACE) for selected parameter sets with sample size n = 10, 20, 40, 100.

TABLE VI

ASYMPTOTIC COMPARISON OF PROP AND ACE WITH Tr(Q)

ρ	ε	Tr(Q)	nxMSE(PROP)	nxMSE(ACE)	Efficiency PROP	Efficiency ACE
.01	1.00	2.037	2.037	3.921	1.000	.519
	.50	1.270	1.270	2.206	1.000	.576
	.05	1.009	1.013	1.081	.996	.933
.05	1.00	2.128	2.136	3.628	.996	.587
	.50	1.317	1.326	2.041	.994	.646
	.10	1.007	1.050	1.098	.959	.918
.10	1.00	2.153	2.176	3.306	.990	.651
	.70	1.595	1.617	2.388	.987	.668
	.50	1.320	1.348	1.860	.978	.710
	.30	1.115	1.162	1.397	.959	.798
.30	1.00	1.849	1.945	2.367	.951	.781
	.70	1.349	1.447	1.710	.932	.789
	.50	1.069	1.210	1.331	.883	.803
.50	1.00	1.448	1.593	1.778	.909	.814
	.70	1.028	1.188	1.284	.865	.800

TABLE VII

SIMULATED COMPARISON OF MLE, PROP, ACE

ρ	ε	Sample Size n	nxMSE(MLE)	nxMSE(PROP)	nxMSE(ACE)	Efficiency MLE	PROP	ACE
.05	1.0	10	3.51	2.68	4.39	.61	.79	.49
		20	2.66	2.33	3.90	.80	.92	.55
		40	2.45	2.27	3.83	.87	.94	.56
		100	2.31	2.24	3.74	.92	.95	.57
.10	1.0	10	3.50	2.71	3.96	.62	.79	.54
		20	2.67	2.35	3.47	.81	.92	.61
		40	2.44	2.29	3.48	.88	.94	.62
		100	2.32	2.27	3.33	.93	.95	.63
		200	2.18	2.20	3.31	.99	.98	.65
.10	.50	10	2.06	1.60	2.14	.64	.82	.62
		20	1.74	1.52	2.03	.76	.87	.65
		40	1.51	1.42	1.97	.88	.93	.67
		100	1.37	1.36	1.90	.96	.97	.70
.30	1.0	10	2.80	2.33	2.75	.66	.79	.67
		20	2.20	2.08	2.49	.84	.88	.73
		40	2.09	2.10	2.53	.89	.89	.74
		100	1.93	1.99	2.44	.96	.93	.76

By comparing the entries of Table VII and those of Table VI it is demonstrated that the method of maximum likelihood estimation and the method PROP has better efficiency than the method ACE.

Table VIII compares the asymptotic mean square error times sample size for individual estimates using the method PROP and ACE with the Cramér-Rao lower bounds. It appears that for ρ large, the estimation of λ_1 and λ_2 by ACE is more efficient than by PROP. However, for other combinations PROP is superior for λ_1 and λ_2. To estimate λ_3 the PROP method is clearly indicated as more efficient.

Table VII gives results of simulations for selected parameter combinations and sample size. One may compare these results with entries in Table VI to have an indication of the behavior of the mean square error times n as n becomes large in relation to Tr(Q).

TABLE VIII

SAMPLE SIZE TIMES ASYMPTOTIC MEAN SQUARE ERROR

FOR INDIVIDUAL ESTIMATES

		Lower Bound			nxMSE(PROP)			nxMSE(ACE)		
		λ_1	λ_2	λ_3	λ_1	λ_2	λ_3	λ_1	λ_2	λ_3
ρ=.01	ε= .05	.010	.990	.009	.012	.990	.011	.041	1.029	.011
	.10	.019	.990	.011	.019	.990	.012	.097	1.077	.012
	.30	.098	.991	.016	.099	.991	.016	.370	1.271	.017
	.50	.257	.992	.022	.257	.992	.022	.721	1.463	.022
	.70	.494	.994	.028	.494	.994	.028	1.150	1.655	.028
	.90	.811	.997	.035	.811	.997	.035	1.658	1.846	.035
	1.00	.999	.999	.038	.999	.999	.038	1.941	1.941	.039
ρ=.05	ε= .10	.028	.950	.029	.050	.950	.050	.050	.992	.055
	.30	.117	.951	.064	.124	.951	.071	.295	1.161	.077
	.50	.270	.956	.091	.275	.956	.095	.612	1.327	.102
	.70	.498	.964	.119	.501	.965	.122	1.002	1.488	.131
	.90	.801	.976	.149	.803	.976	.153	1.465	1.646	.164
	1.00	.981	.981	.166	.983	.983	.169	1.723	1.723	.181
ρ=.10	ε= .30	.117	.900	.097	.141	.900	.121	.215	1.042	.140
	.50	.266	.905	.149	.280	.906	.163	.496	1.178	.186
	.70	.481	.914	.200	.491	.916	.210	.843	1.307	.239
	.90	.767	.928	.253	.774	.932	.263	1.257	1.429	.299
	1.00	.936	.936	.280	.942	.942	.291	1.488	1.488	.330
ρ=.30	ε= .50	.146	.687	.236	.214	.698	.299	.178	.754	.399
	.70	.318	.678	.353	.363	.692	.392	.402	.795	.513
	.90	.538	.671	.461	.572	.694	.493	.675	.821	.641
	1.00	.668	.668	.513	.699	.699	.547	.828	.828	.710
ρ=.50	ε= .70	.140	.466	.423	.211	.494	.483	.151	.491	.642
	.90	.308	.428	.570	.362	.468	.612	.338	.464	.802
	1.00	.406	.406	.635	.457	.457	.680	.444	.444	.889
ρ=.70	ε= .90	.128	.238	.676	.180	.279	.718	.131	.243	.874
	1.00	.202	.202	.756	.251	.251	.797	.208	.208	.969
ρ=.90	ε=1.00	.055	.055	.907	.078	.078	.927	.055	.055	.997

VI. TEST FOR DEPENDENCE

A. Introduction

The main problem considered in this chapter is constructing a test for the hypothesis that the correlation ρ between X and Y is zero against the alternative that ρ is positive when $(X,Y) \sim BVE(\lambda_1, \lambda_2, \lambda_3)$. The parameter λ_3 will be allowed to take on the value zero since $P[X=Y] = \rho = \lambda_3/\lambda$; that is, $\rho=0$ if and only if $\lambda_3=0$. A simple test based on the statistic n_3 = number of sample points (X_j, Y_j) such that $X_j=Y_j$ is given first which does not depend on knowledge of any of the parameters. An indication of how this test may be extended to the multivariate exponential distribution is also given. A test for $H:\rho=0$ against $K:\rho=\rho_o>0$ is then considered for the case when the marginal parameters $\gamma_1=\lambda_1+\lambda_3$, $\gamma_2=\lambda_2+\lambda_3$ are known. This test is uniformly most powerful for $H:\rho=0$ against the alternative $K:\rho>0$ when $\gamma_1=\gamma_2$. Then a test of $H:\rho=0$ against $K:\rho>0$ is given for the case when the marginal parameters are equal $(\gamma_1=\gamma_2)$ but unknown. Finally the matter of constructing confidence intervals for ρ and the marginal parameters is considered.

The test for $\rho=0$ against $\rho>0$ is made to determine whether the bivariate exponential distribution with $\lambda_3>0$ is the appropriate model or whether the assumption of independent exponentials is appropriate. Some errors associated with assuming independence when in fact the random variables X and Y are positively correlated are illustrated in Chapter VII. Instead of using λ_1, λ_2, λ_3 as the parameters for the BVED, the parameters γ_1, γ_2, ρ will be used. The reason for this will become clear in Chapter VII.

The following relations are clear if $\gamma_1 = \lambda_1+\lambda_3$, $\gamma_2 = \lambda_2+\lambda_3$.

$$
\begin{aligned}
\lambda_1 &= (\gamma_1-\rho\gamma_2)/(1+\rho) \\
\lambda_2 &= (\gamma_2-\rho\gamma_1)/(1+\rho) \\
\lambda_3 &= \rho(\gamma_1+\gamma_2)/(1+\rho) \\
\lambda &= (\gamma_1+\gamma_2)/(1+\rho)
\end{aligned}
\tag{9}
$$

If $\rho=0$ then it is seen that $\gamma_1=\lambda_1$, $\gamma_2=\lambda_2$, $\lambda_3=0$ and $\lambda=\gamma_1+\gamma_2$. The bivariate exponential distribution which has been denoted $BVE(\lambda_1, \lambda_2, \lambda_3)$ will also now be denoted by $BVE(\gamma_1, \gamma_2, \rho)$ where the context of the notation will clearly denote which set of parameters is being used, whether $(\lambda_1, \lambda_2, \lambda_3)$ or $(\gamma_1, \gamma_2, \rho)$. If $\rho=0$ then $BVE(\gamma_1, \gamma_2, \rho)$ is just the product of two independent exponential distributions with parameters γ_1 and γ_2, respectively.

B. A Test Based on n_3

Suppose $\{(X_j, Y_j)\}$ is a random sample of size n from $BVE(\gamma_1,\gamma_2,\rho)$ $(\gamma_1>0, \gamma_2>0, \rho\geq 0)$. Since $P[X=Y] = \rho$ it is natural to use as a test statistic n_3 which counts the number of sample points (X_j, Y_j) for which $X_j=Y_j$. Clearly if $n_3>0$ then $\rho>0$. Thus, to test $H:\rho=0$ against the alternative $K:\rho>0$ one could reject H if $n_3>0$. The probability of type I error is zero; that is, $P[n_3>0|H] = 0$. For when H is true the probability that X=Y is zero. The power of the test is $P[n_3>0|K] = 1 - P[n_3=0|K]$. Since n_3 has a binomial distribution with parameters n and ρ, $P[n_3=0|K] = (1-\rho)^n$. Hence the power of this test is $1 - (1-\rho)^n$. The power for various ρ and sample sizes n is given in Table IX.

As an indication of how this test may be extended to the multivariate exponential distribution given in [1], consider a particular

TABLE IX

POWER OF n_3 TEST

Sample Size

ρ	2	4	6	8	10	15
.01	.0199	.0394	.0585	.0773	.0956	.1399
.03	.0591	.1147	.1670	.2163	.2626	.3677
.05	.0975	.1855	.2649	.3366	.4013	.5367
.07	.1351	.2519	.3530	.4404	.5160	.6633
.10	.1900	.3439	.4686	.5695	.6513	.7941
.20	.3600	.5904	.7379	.8322	.8926	.9648
.30	.5100	.7599	.8824	.9424	.9718	.9953
.40	.6400	.8704	.9533	.9832	.9940	.9995
.50	.7500	.9375	.9844	.9961	.9990	1.0000
.60	.8400	.9744	.9959	.9993	.9999	1.0000

Sample Size

ρ	20	25	30	40	50	100
.01	.1821	.2222	.2603	.3310	.3950	.6340
.03	.4562	.5330	.5990	.7043	.7819	.9524
.05	.6415	.7226	.7854	.8715	.9231	.9941
.07	.7658	.8370	.8866	.9451	.9734	.9993
.10	.8784	.9282	.9576	.9852	.9948	1.0000
.20	.9885	.9962	.9988	.9999	1.0000	1.0000
.30	.9992	.9999	1.0000	1.0000	1.0000	1.0000
.40	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
.50	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
.60	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

trivariate exponential distribution with only three parameters λ_1, λ_2, λ_3 denoted by TVE(λ_1, λ_2, λ_3); where it is assumed that X, Y, Z are random variables which take on nonnegative values and $\overline{F}(x,y,z;\lambda_1,\lambda_2,\lambda_3) = Pr[X>x, Y>y, Z>z] = \exp\{-\lambda_1(x+y+z) - \lambda_2[\max(x,y) + \max(x,z) + \max(y,z)] - \lambda_3[\max(x,y,z)]\}$ where $\lambda_1>0$ and λ_2 and λ_3 are nonnegative. It is found that $P[X=Y=Z] = \lambda_3/(3\lambda_1+3\lambda_2+\lambda_3)$ and $P[X=Y] = P[X=Z] = P[Y=Z] = (\lambda_1+\lambda_2)(3\lambda_2+\lambda_3)/(2\lambda_1+3\lambda_2+\lambda_3)(3\lambda_1+3\lambda_2+\lambda_3)$.

Given a random sample of size n from TVE(λ_1, λ_2, λ_3) let n_1 count the number of (X_j, Y_j, Z_j) such that no ties occur; that is the number of j such that $X_j \neq Y_j$ and $Y_j \neq Z_j$ and $X_j \neq Z_j$. Let n_2 count the number of (X_j, Y_j, Z_j) such that $X_j = Y_j \neq Z_j$ or $X_j=Z_j \neq Y_j$ or $Y_j=Z_j \neq X_j$; that is, the number of pairwise ties. Let n_3 count the number of (X_j, Y_j, Z_j) such that a threeway tie occurs, $X_j=Y_j=Z_j$. Note that $n_1+n_2+n_3=n$. To test $H:\lambda_2=\lambda_3=0$ against the alternative $K:\lambda_2>0$ or $\lambda_3>0$ based on a random sample of size n from TVE(λ_1, λ_2, λ_3) reject H if $n_2+n_3>0$.

Let $\beta_3 = P[X=Y=Z]$, $\beta_2 = P[X=Y\neq Z \text{ or } X=Z\neq Y \text{ or } Y=Z\neq X]$ and $\beta_1 = 1-\beta_2-\beta_3$. Then ($n_1$, n_2) is distributed as a trinomial distribution with parameters n and β_1 and β_2. The probability that $n_1 = k_1$ and $n_2 = k_2$ (hence $n_3 = n-k_1-k_2 = k_3$) is given by

$$n!\,\beta_1^{k_1}\ \beta_2^{k_2}\ \beta_3^{k_3}/k_1!k_2!k_3!.$$

Hence the probability of rejecting H, when H is true, is zero since $\lambda_2=\lambda_3=0$ imply $\beta_2=\beta_3=0$. The power of the test $P[n_2+n_3>0|K] = 1 - P[n_1=0,\ n_2=0|K] = 1-\beta_1^n$. For example, if $\lambda_1 = 1$, $\lambda_2 = .1$, $\lambda_3 = .1$ and n=10 the power of the test is .880.

In a life test situation it may be that only $\min(X_j, Y_j, Z_j)$ is observed for each j. In this case the event $X_j=Y_j$ would not be observable if $Z_j<X_j$. One may consider, however, the following events. Let A be the event that $X<\min(Y,Z)$ or $Y<\min(X,Z)$ or $Z<\min(X,Y)$; let B be the event that $X=Y<Z$ or $X=Z<Y$ or $Y=Z<X$; let C be the event $X=Y=Z$. It is found that $P[C] = \lambda_3/(3\lambda_1+3\lambda_2+\lambda_3)$ as before, but $P[B] = 3\lambda_2/(3\lambda_1+3\lambda_2+\lambda_3)$ and $P[A] = 1 - P[B] - P[C]$. If m_1, m_2, m_3 count the number of times events A, B, C occur, respectively, then reject $H:\lambda_2=\lambda_3=0$ in favor of $K:\lambda_2>0$ or $\lambda_3>0$ if $m_2+m_3>0$. If $\beta_1^* = P[A]$ then the probability of a type I error is zero as before but the power is $1-(\beta_1^*)^n$. For example, if $\lambda_1=1$, $\lambda_2=.1$, $\lambda_3=.1$ and $n=10$ the power of the test is .714.

Again this difference in power from .880 to .714 indicates that if all observations are not available there is a loss of information, in this case a loss of power. The extension of this procedure to the multivariate exponential distribution is clear based on the multinomial distribution of the counting statistics.

C. <u>A Test for the Simple Hypothesis and Simple Alternative</u>.

To test the simple hypothesis $H:\rho=0$, γ_1, γ_2 known against the simple alternative $K:\rho = \rho_o>0$, γ_1, γ_2 known at level α given a random sample of size n from BVE(γ_1, γ_2, ρ), the method of the Neyman-Pearson lemma [15] requires that H be rejected when the ratio of the likelihood under K to the likelihood under H is greater than some constant chosen so that the probability of rejection when H is true is α. Let L_H and L_K denote the likelihood under H and K, respectively. So H is rejected when $L_H/L_K>\theta_o$ where $P[L_K/L_H > \theta_o|H] = \alpha$. The power of the test is $P[L_K/L_H > \theta_o|K]$. Equivalently H is rejected when $\ln(L_K/L_H) = \phi>\phi_o = \ln\theta_o$. The likelihood of the sample is defined (see page 351 in Wilks [10]) to be $L = \Pi dF(X_j, Y_j, \gamma_1, \gamma_2, \rho)$ where

$$dF(X,Y;\ \gamma_1,\ \gamma_2,\ 0) = \gamma_1\gamma_2\exp\{-\gamma_1X-\gamma_2Y\}hk$$

$$dF(X,Y;\ \gamma_1,\ \gamma_2,\ \rho_o) = \begin{cases} \lambda_1\gamma_2\exp\{-\lambda_1X-\gamma_2Y\}hk & \text{if } X<Y \\ \lambda_2\gamma_1\exp\{-\gamma_1X-\lambda_2Y\}hk & \text{if } X>Y \\ \exp\{-\lambda X\}[\exp\{-\lambda_1h-\lambda_2k-\lambda_3\max(h,k)\} + 1 & \\ \qquad - \exp\{-\gamma_1h\} - \exp\{-\gamma_2k\}] & \text{if } X=Y \end{cases} \qquad (10)$$

$$(h>0,\ k>0)$$

In (10) the values λ_1, λ_2, λ_3, λ are to be interpreted to mean those values obtained through relations (9) with the known values of γ_1, γ_2 and ρ_o inserted.

With the remark above in mind the expressions below follow where n_1 counts the number of $(X_j,\ Y_j)$ such that $X_j<Y_j$, n_2 counts the number of $(X_j,\ Y_j)$ such that $X_j>Y_j$ and n_3 counts the number of $(X_j,\ Y_j)$ such that that $X_j=Y_j$.

$$L_H = \gamma_1{}^n\gamma_2{}^n\exp\{-\gamma_1\sum X_j\ -\gamma_2\sum Y_j\}h^nk^n$$

$$L_K = \lambda_1{}^{n_1}\gamma_2{}^{n_1}\lambda_2{}^{n_2}\gamma_1{}^{n_2}\exp\{-\gamma_1\sum X_j\ -\ \gamma_2\sum Y_j\ +\ \lambda_3\sum\min(X_j,Y_j)\}h^{n_1+n_2}k^{n_1+n_2}$$

$$\cdot[\exp\{-\lambda_1h-\lambda_2k-\lambda_3\max(h,k)\} + 1 - \exp\{-\gamma_1h\} - \exp\{-\gamma_2k\}]^{n_3}$$

Now $\phi = \ln(L_K/L_H)$ simplifies to the expression below where AU = $\sum\min(X_j,\ Y_j)$.

$$\phi = n_1\ln(\lambda_1/\gamma_1) + n_2\ln(\lambda_2/\gamma_2) + \lambda_3AU - \qquad (11)$$

$$n_3(\ln[\gamma_1\gamma_2hk]-\ln[\exp\{-\gamma_1h-\gamma_2k+\lambda_3\min(h,k)\}+1-\exp\{-\gamma_1h\}-\exp\{-\gamma_2k\}])$$

1. $\gamma_1 = \gamma_2$ Known

Suppose now that γ_1 and γ_2 are known and equal. Let γ denote their common value. Then (11) simplifies to the following:

$$\phi = (n_1+n_2)\ \ln[(1-\rho_o)/(1+\rho_o)] + 2\gamma\rho_o AU/(1+\rho_o) -$$
$$n_3\ [\ln(\gamma^2 hk) - \ln\{\exp[-\gamma(h+k) + 2\gamma\rho_o \min(h,k)/(1+\rho_o)]$$
$$+ 1 - \exp(-\gamma h) - \exp(-\gamma k)\}]. \tag{12}$$

Under H, $n_3=0$ so $P[\phi \geq \phi_o | H] = P[n\ \ln(1-\rho_o)/(1+\rho_o) + 2\gamma\rho_o AU/(1+\rho_o) \geq \phi_o | H]$ $= P[(2\gamma\rho_o AU)/(1+\rho_o) \geq \phi_o + n\ \ln(1+\rho_o)/(1-\rho_o) | H]$. Now AU has a gamma distribution with parameters n and $1/\lambda$, so $2\lambda AU$ is distributed as a Chi-Square random variable with 2n degrees of freedom. Under H, λ is 2γ, whereas under K, λ is $2\gamma/(1+\rho_o)$. Hence $P[\phi \geq \phi_o | H] =$

$$P[4\gamma AU \geq 2(1+\rho_o)\{\phi_o + n\ \ln(1+\rho_o)/(1-\rho_o)\}/\rho_o | H] =$$
$$P[\chi^2(2n) \geq 2(1+\rho_o)\{\phi_o + n\ \ln(1+\rho_o)/(1-\rho_o)\}/\rho_o].$$

So choosing $\phi_o = \rho_o\chi^2_\alpha(2n)/2(1+\rho_o) + n\ \ln(1-\rho_o)/(1+\rho_o)$ will give an α-level test for H against K where $\chi^2_\alpha(2n)=c$ is found such that $\int_c^\infty f(u)du=\alpha$ where f(u) is the Chi-Square density with parameter 2n.

The power of the test is found by computing $P[\phi \geq \phi_o | K]$. Let $P_K(E)$ denote the probability of event E when K is true. Consider

$P_K[\phi \geq \phi_o] = P_K[n_3=0]P_K[\phi \geq \phi_o | n_3=0] + P_K[n_3 \geq 1]P_K[\phi \geq \phi_o | n_3>0]$. Now

$P_K[n_3=0] = (1-\rho_o)^n$ and $P_K[n_3>0] = 1 - (1-\rho_o)^n$. Also $P_K[\phi \geq \phi_o | n_3=0]$

$= P_K[2\gamma\rho_o AU/(1+\rho_o) \geq \rho_o\chi^2_\alpha(2n)/2(1+\rho_o) | n_3=0] = P[\chi^2(2n) \geq \chi^2_\alpha(2n)/(1+\rho_o)]$

since $4\gamma/(1+\rho_o)AU \sim \chi^2(2n)$ under K.

It will be shown now that $P_K[\phi \geq \phi_o | n_3>0] = 1$. Consider the expression (12) as $h\to 0^+$, $k\to 0^+$. The first two terms in (12) are not dependent on h or k. Consider the coefficient of $-n_3$ repeated here:

$$\ln(\gamma^2 hk) - \ln[\exp\{-\gamma(h+k)+(2\gamma\rho_o \min(h,k)/(1+\rho_o))\} + 1 - \exp\{-\gamma h\} - \exp\{-\gamma k\}].$$

The quantity in the square brackets above is $2\gamma\rho_o \min(h,k)/(1+\rho_o) + D_2$, where D_2 denotes terms of degree 2 or more in h and k. So $\ln(\gamma^2 hk) - \ln[\;] = \ln\{\gamma^2 hk(1+\rho_o)/[2\gamma\rho_o \min(h,k)+D_2]\}$. As $h\to 0^+$, $k\to 0^+$ the argument of the logarithm tends to zero so that the coefficient of $-n_3$ tends to $-\infty$. If $n_3>0$ then it is clear that ϕ given in (12) exceeds any constant ϕ_o, or in other words $P_K[\phi \geq \phi_o | n_3>0] = 1$.

In summary then the power of the test of $H:\rho=0$, $\gamma_1=\gamma_2$ known against the alternative $K:\rho=\rho_o>0$, $\gamma_1=\gamma_2$ known is

$$P_K[\phi \geq \phi_o] = 1 - (1-\rho_o)^n + (1-\rho_o)^n \, P[\chi^2(2n) \geq \chi^2_\alpha(2n)/(1+\rho_o)].$$

The first term on the right is just the power of the n_3 test given in section B. Table X gives the ϕ_o values and Table XI gives the power of the test for various ρ_o and sample sizes and $\alpha=.05, .10, .20$. In practice the procedure is to reject H if $n_3>0$ or if $n_3=0$ and $\sum \min(X_j, Y_j) \geq \chi^2_\alpha(2n)/4\gamma$ where γ is the common known value of γ_1 and γ_2. This gives an α-level test of H against K. It is clear that this test is also uniformly most powerful for $H:\rho=0$, $\gamma_1=\gamma_2=\gamma$ known against $K:\rho>0$, $\gamma_1=\gamma_2=\gamma$ known.

2. $\gamma_1 \neq \gamma_2$ Known

Suppose γ_1 and γ_2 are both known but $\gamma_1 \neq \gamma_2$. Let γ denote $\max(\gamma_1, \gamma_2)$ and $\varepsilon=\min(\gamma_1, \gamma_2)/\gamma$. Note that $0<\varepsilon<1$. Without loss of generality assume $\gamma_1=\varepsilon\gamma$ and $\gamma_2=\gamma$; that is, γ_2 is $\max(\gamma_1, \gamma_2)$. Then in terms of γ, ε and ρ_o, the expression (11) becomes

$$\phi = n_1 \ln\{(\varepsilon-\rho_o)/\varepsilon(1+\rho_o) \quad + n_2 \ln\{(1-\varepsilon\rho_o)/(1+\rho_o)\} + \gamma\rho_o(\varepsilon+1)AU/(1+\rho_o)$$

$$- n_3\big(\ln[\varepsilon\gamma^2 hk] + 1 - \exp\{-\gamma h\} - \exp\{-\gamma k\}$$

$$-\ln[\exp\{-\varepsilon\gamma h-\gamma k+\gamma\rho_o(\varepsilon+1)\min(h,k)/(1+\rho_o)]\big)$$

Under $H:\rho=0$, $n_3=0$ so $P_H[\phi \geq \phi_o] =$

$$P_H[n_1 \ln\{(\varepsilon-\rho_o)/\varepsilon(1+\rho_o)\} + n_2 \ln\{(1-\varepsilon\rho_o)/(1+\rho_o)\} + [\gamma\rho_o(\varepsilon+1)AU/(1+\rho_o) \geq \phi_o].$$

Now $P_H[\phi \geq \phi_o] = \sum_{n_1=0}^{n} P_H[n_1]P_H[\phi \geq \phi_o|n_1] = \sum_{n_1=0}^{n} n!\varepsilon^{n_1}/n_1!n_2!(1+\varepsilon)^n \cdot P_H[\phi \geq \phi_o|n_1]$

since n_1 has a binomial distribution with parameters n and $\varepsilon/(\varepsilon+1)$ under H and $n_1+n_2=n$. $P_H[\phi \geq \phi_o|n_1] =$

$$P_H[\gamma\rho_o(\varepsilon+1)AU/(1+\rho_o) \geq \phi_o + n_1 \ln\{\varepsilon(1+\rho_o)/(\varepsilon-\rho_o)\} + n_2 \ln\{(1+\rho_o)/(1-\varepsilon\rho_o)\}|n_1]$$

$$= P[\chi^2(2n) \geq 2(1+\rho_o)\big(\phi_o + n_1 \ln\{\varepsilon(1+\rho_o)/(\varepsilon-\rho_o)\} + n_2 \ln\{(1+\rho_o)/(1-\varepsilon\rho_o)\}/\rho_o\big)$$

For fixed n_1 (and n_2) as ϕ_o increases the probability above decreases so that ϕ_o can be chosen so that the above probability takes on a given value. In particular ϕ_o can be chosen so that

$$P_H[\phi \geq \phi_o] = \sum_{n_1=0}^{n} P_H[\phi \geq \phi_o|n_1]P_H(n_1)$$

takes on a given value α which is the probability of the type I error. Given the values n, γ, ε, ρ_o and α the value of ϕ_o may be determined. Table XI gives ϕ_o for various n, ε, ρ_o and α = .05, .10, .20. The maximum marginal parameter γ is taken as 1 since when γ_1 and γ_2 are known the scale may be normalized with respect to the maximum of γ_1 and γ_2.

To compute the power of the test consider

$$P_K[\phi \geq \phi_o] = P_K[n_3=0]P_K[\phi \geq \phi_o|n_3=0] + P_K[n_3>0]\ P_K[\phi \geq \phi_o|n_3>0].$$

As in Section C-1

$$P_K[n_3>0] = 1 - (1-\rho_o)^n$$

$$P_K[n_3=0] = (1-\rho_o)^n$$

$$P_K[\phi \geq \phi_o|n_3>0] = 1.$$

The argument for the last equality is similar to that in Section C-1, so will not be repeated here. To compute $P_K[\phi \geq \phi_o|n_3=0]$ consider further conditioning on n_1.

$$P_K[\phi \geq \phi_o|n_3=0] = \sum_{n_1=0}^{n} P_K[n_1|n_3=0]\ P_K[\phi \geq \phi_o|n_1|n_3=0].$$

$$P_K[n_1|n_3=0] = n!(\varepsilon-\rho_o)^{n_1}(1-\rho_o\varepsilon)^{n_2}/n_1!n_2!(1+\varepsilon)^n(1-\rho_o)^n \quad (13)$$

$P_K[\phi \geq \phi_o|n_1|n_3=0] =$

$$P_K\left(2\gamma(1+\varepsilon)AU/(1+\rho_o) \geq 2\{\phi_o+n_1\ln[\varepsilon(1+\rho_o)/(\varepsilon-\rho_o)]+n_2\ln[(1+\rho_o)/(1-\varepsilon\rho_o)]\}/\rho_o\right)$$

As noted before $2\lambda AU \sim \chi^2(2n)$ where under K $\lambda=\gamma(1+\varepsilon)/(1+\rho_o)$. Hence

$P_K[\phi \geq \phi_o|n_1|n_3=0] =$

$$P\left(\chi^2(2n) \geq 2\{\phi_o + n_1\ln[\varepsilon(1+\rho_o)/(\varepsilon-\rho_o)] + n_2\ln[(1+\rho_o)/(1-\varepsilon\rho_o)]\}/\rho_o\right) \quad (14)$$

Putting together (13) and (14) and summing over $n_1=0, 1,\ldots, n$ with $n_2=n-n_1$ the power $P_K[\phi \geq \phi_o]$ may be computed. The power of the test is

$$[1 - (1-\rho_o)^n] + \sum_{n_1=0}^{n} n!(\varepsilon-\rho_o)^{n_1}(1-\varepsilon\rho_o)^{n_2}/n_1!n_2!(1+\varepsilon)^n \cdot$$

$$P\left(\chi^2(2n) \geq 2\{\phi_o + n_1\ln[\varepsilon(1+\rho_o)/(\varepsilon-\rho_o)] + n_2\ln[(1+\rho_o)/(1-\varepsilon\rho_o)]\}/\rho_o\right)$$

Again note that the first term is just the power from the n_3 test given in B. Table X gives the power of various n, ε, ρ_o and α = .05, .10, .20.

In practice reject $H:\rho=0$, $\gamma_1 \neq \gamma_2$ known in favor of $K:\rho=\rho_o>0$, $\gamma_1 \neq \gamma_2$ known if $n_3>0$ or if $n_3=0$ and $\sum \min(X_j, Y_j) \geq$

$$(1+\rho_o)\{\phi_o + n_1 \ln[\varepsilon(1+\rho_o)/(\varepsilon-\rho_o)] + n_2 \ln[(1+\rho_o)/(1-\varepsilon\rho_o)]\}/\gamma\rho_o(1+\varepsilon)$$

where ϕ_o is given in Table XI.

If $\varepsilon=1$, equivalently if $\gamma_1=\gamma_2$, then the test given in this section yields the same test as in Section C-1. When using Table X or XI the entries marked $\varepsilon=1$ correspond to the test in Section C-1 where it is assumed that the common value γ is scaled to unity. The test developed here in Section C will be referred to as the AU-n_3 test.

D. $\gamma_1 = \gamma_2$ Unknown

Suppose both γ_1 and γ_2 are unknown but it is known that $\gamma_1=\gamma_2$. Let γ denote the common unknown value for γ_1 and γ_2. To test the composite hypothesis $H:\rho=0$, $\gamma_1=\gamma_2$ unknown against the composite alternative $K:\rho>0$, $\gamma_1=\gamma_2$ unknown the following procedure is suggested. Given a random sample of size n from BVE(γ, γ, ρ) reject H if $n_3>0$ or if $n_3=0$ and the statistic $4AU/(AX+AY)>\phi_o$ where $AU = \sum \min(X_j,Y_j)$, $AX = \sum X_j$, $AY = \sum Y_j$. Here ϕ_o is a function of the sample size n and level α of the test and may be found for given n and α by simulating the distribution of $4AU/(AX+AY)$ with $\rho=0$. ϕ_o then cuts off the upper α tail. The test appears to be independent of γ. Table XII below gives ϕ_o for n = 10, 20, 40 and α = .05, .10, .20. The power for this test has been found by simulation for various ρ values. These are given in Table XIII for sample sizes n=10, 20, 40 and α=.20, .10, .05.

TABLE X

POWER OF THE AU-n_3 TEST

α = .20

		Sample Size				
ρ	ε	4	10	20	40	100
.01	1.0	.238	.285	.357	.477	.718
	.9	.238	.285	.357	.478	.718
	.7	.238	.286	.357	.478	.719
	.5	.239	.287	.359	.480	.721
	.3	.241	.290	.363	.485	.725
	.1	.243	.298	.377	.502	.742
	.05	.245	.302	.387	.520	.760
.05	1.0	.374	.551	.738	.910	.996
	.9	.374	.551	.738	.910	.996
	.7	.375	.552	.740	.911	.996
	.5	.379	.557	.745	.914	.997
	.3	.388	.570	.756	.920	.997
	.1	.406	.613	.808	.949	.999
.10	1.0	.517	.756	.920	.991	
	.9	.517	.756	.921	.991	
	.7	.519	.759	.922	.992	
	.5	.527	.766	.926	.992	
	.3	.545	.786	.937	.994	
.30	1.0	.853	.986	1.000		
	.9	.854	.986	1.000		
	.7	.860	.987	1.000		
	.5	.879	.991	1.000		
.50	1.0	.969	1.000			
	.9	.969	1.000			
	.7	.975	1.000			

TABLE X Continued

$\alpha = .10$

ρ	ε	Sample Size 4	10	20	40	100
.01	1.0	.140	.192	.271	.406	.678
	.9	.140	.192	.271	.406	.678
	.7	.140	.192	.272	.407	.678
	.5	.140	.193	.272	.408	.679
	.3	.141	.195	.275	.411	.682
	.1	.143	.199	.283	.422	.693
	.05	.143	.200	.287	.432	.706
.05	1.0	.285	.481	.695	.893	.995
	.9	.285	.481	.695	.893	.995
	.7	.285	.482	.696	.894	.995
	.5	.288	.486	.699	.896	.996
	.3	.292	.493	.707	.901	.996
	.1	.302	.517	.742	.924	.998
.10	1.0	.439	.711	.903	.989	
	.9	.439	.711	.903	.989	
	.7	.441	.713	.904	.989	
	.5	.445	.718	.907	.990	
	.3	.456	.731	.916	.992	
.30	1.0	.819	.982	1.000		
	.9	.819	.982	1.000		
	.7	.824	.983	1.000		
	.5	.838	.986	1.000		
.50	1.0	.959	1.000			
	.9	.960	1.000			
	.7	.965	1.000			

TABLE X Continued

$\alpha = .05$

ρ	ε	Sample Size 4	10	20	40	100
.01	1.0	.090	.144	.227	.370	.657
	.9	.090	.144	.227	.370	.657
	.7	.090	.145	.228	.370	.657
	.5	.091	.145	.228	.370	.658
	.3	.091	.146	.230	.372	.659
	.1	.092	.148	.234	.378	.666
	.05	.092	.149	.236	.383	.674
.05	1.0	.238	.444	.670	.884	.995
	.9	.238	.444	.670	.884	.995
	.7	.238	.445	.671	.884	.995
	.5	.239	.446	.673	.885	.995
	.3	.242	.451	.678	.888	.995
	.1	.246	.463	.698	.905	.997
.10	1.0	.396	.685	.893	.988	
	.9	.396	.685	.893	.988	
	.7	.397	.686	.894	.988	
	.5	.400	.689	.896	.988	
	.3	.406	.698	.902	.990	
.30	1.0	.797	.978	1.000		
	.9	.797	.979	1.000		
	.7	.800	.979	1.000		
	.5	.809	.982	1.000		
.50	1.0	.953	.999			
	.9	.953	.999			
	.7	.957	1.000			

TABLE XI

CRITICAL VALUES FOR THE AU-n_3 TEST

$\alpha = .20$

		Sample Size				
ρ	ε	4	10	20	40	100
.01	1.0	-.025	-.076	-.166	-.352	-.928
	.9	-.025	-.076	-.166	-.352	-.927
	.7	-.024	-.074	-.164	-.349	-.923
	.5	-.021	-.070	-.157	-.341	-.910
	.3	-.014	-.058	-.142	-.320	-.881
	.1	.004	-.015	-.088	-.249	-.784
	.05	.009	.006	-.018	-.194	-.698
.05	1.0	-.138	-.405	-.876	-1.851	-4.851
	.9	-.137	-.404	-.875	-1.850	-4.850
	.7	-.132	-.396	-.866	-1.838	-4.838
	.5	-.116	-.375	-.839	-1.807	-4.809
	.3	-.077	-.322	-.776	-1.741	-4.770
	.1	.014	-.084	-.615	-1.637	-5.133
.10	1.0	-.301	-.869	-1.865	-3.918	
	.9	-.300	-.867	-1.863	-3.916	
	.7	-.289	-.853	-1.848	-3.903	
	.5	-.257	-.813	-1.808	-3.876	
	.3	-.174	-.723	-1.736	-3.880	
.30	1.0	-1.204	-3.301	-6.927		
	.9	-1.200	-3.299	-6.930		
	.7	-1.161	-3.285	-6.979		
	.5	-1.056	-3.322	-7.309		
.50	1.0	-2.556	-6.813			
	.9	-2.548	-6.822			
	.7	-2.494	-7.030			

TABLE XI Continued

α = .10

ρ	ε	4	10	20	40	100
		Sample Size				
.01	1.0	-.014	-.059	-.144	-.322	-.881
	.9	-.014	-.059	-.143	-.322	-.880
	.7	-.013	-.057	-.140	-.317	-.874
	.5	-.009	-.051	-.132	-.305	-.855
	.3	-.001	-.037	-.111	-.275	-.809
	.1	.016	.008	-.036	-.174	-.657
	.05	.021	.025	.015	-.056	-.514
.05	1.0	-.082	-.324	-.768	-1.704	-4.627
	.9	-.082	-.323	-.767	-1.702	-4.624
	.7	-.075	-.313	-.752	-1.683	-4.600
	.5	-.056	-.283	-.712	-1.630	-4.534
	.3	-.013	-.210	-.615	-1.510	-4.401
	.1	.073	.024	-.222	-1.149	-4.322
.10	1.0	-.195	-.715	-1.659	-3.637	
	.9	-.194	-.713	-1.656	-3.634	
	.7	-.180	-.692	-1.629	-3.604	
	.5	-.140	-.633	-1.557	-3.526	
	.3	-.047	-.489	-1.396	-3.390	
.30	1.0	-.934	-2.912	-6.403		
	.9	-.929	-2.906	-6.400		
	.7	-.872	-2.851	-6.383		
	.5	-.714	-2.744	-6.384		
.50	1.0	-2.168	-6.251			
	.9	-2.153	-6.246			
	.7	-2.020	-6.257			

TABLE XI Continued

α = .05

ρ	ε	Sample Size 4	10	20	40	100
.01	1.0	-.003	-.045	-.124	-.296	-.842
	.9	-.003	-.044	-.124	-.295	-.841
	.7	-.002	-.042	-.120	-.290	-.832
	.5	.002	-.035	-.110	-.275	-.808
	.3	.011	-.018	-.085	-.238	-.750
	.1	.027	.025	-.003	-.116	-.556
	.05	.032	.041	.039	-.001	-.377
.05	1.0	-.031	-.253	-.674	-1.578	-4.437
	.9	-.030	-.252	-.672	-1.575	-4.399
	.7	-.023	-.239	-.655	-1.551	-4.399
	.5	-.002	-.204	-.604	-1.482	-4.304
	.3	.043	-.119	-.482	-1.320	-4.098
	.1	.126	.110	-.046	-.798	-3.683
.10	1.0	-.098	-.579	-1.479	-3.396	
	.9	-.096	-.576	-1.441	-3.392	
	.7	-.080	-.551	-1.441	-3.348	
	.5	-.036	-.478	-1.344	-3.232	
	.3	.062	-.303	-1.120	-2.991	
.30	1.0	-.687	-2.566	-5.947		
	.9	-.680	-2.558	-5.940		
	.7	-.613	-2.474	-5.873		
	.5	-.427	-2.269	-5.806		
.50	1.0	-1.810	-5.751			
	.9	-1.791	-5.736			
	.7	-1.614	-5.645			

TABLE XII

CRITICAL VALUES FOR $H:\rho=0$ VS. $K:\rho>0$
WHEN $\gamma_1=\gamma_2$ UNKNOWN

Sample Size	$\alpha=.20$	$\alpha=.10$	$\alpha=.05$
10	1.191	1.290	1.364
20	1.134	1.204	1.259
40	1.097	1.142	1.180

TABLE XIII

POWER FOR 4AU/(AX+AY) TEST

γ	ρ	n	$\alpha=.20$	$\alpha=.10$	$\alpha=.05$	n_3 Test
1.0	.05	10	.526	.459	.425	.390
		20	.737	.699	.678	.656
		40	.908	.897	.891	.886
		100	.996	.996	.993	.993
1.0	.10	10	.733	.689	.669	.644
		20	.910	.895	.886	.878
		40	.991	.988	.987	.985
1.0	.30	10	.981	.977	.974	.970
		20	.9998	.9998	.9995	.9993
		40	1.0000	1.0000	1.0000	1.0000

The simulated values given in Tables XII and XIII were based on 8,000 samples of size 10, 4000 samples of size 20 and 2000 samples of size 40.

By comparing the power of this test with that of the n_3 test of Section B tabled in Table IX it is found that this test is superior to that of the n_3 test in the case where γ_1 and γ_2 are unknown but equal. The AU-n_3 test of Section C-1 gives an upper bound to the power attainable.

The motivation for the statistic 4AU/(AX+AY) is based on the fact that $2\lambda AU \sim \chi^2(2n)$ and both $2\gamma AX$ and $2\gamma AY$ are distributed as $\chi^2(2n)$. If $\rho=0$ then $2\gamma AX$ and $2\gamma AY$ are independent so that $2\gamma AX+2\gamma AY$ would be distributed as $\chi^2(4n)$. Also if $\rho=0$, $\lambda=2\gamma$. Though AU is not independent of AX+AY, a ratio of ChiSquare statistics divided by their degrees of freedom leads one to consider $(2\lambda AU/2n)/(2\gamma AX+2\gamma AY)/4n = 4AU/(AX+AY)$ with the unknown parameter γ dropping out. This statistic does a reasonable job of indicating whether $\rho=0$ or $\rho>0$ as indicated by Table XIII.

E. Discussion

A test for $H:\rho=0$ against the alternative $K:\rho>0$ has been given in Section B which is valid whether the marginal parameters γ_1, γ_2 are known or not. When the parameters are unknown but equal an improved test is given in Section D. When the parameters γ_1, γ_2 are known and equal then the test given in Section C-2 is uniformly most powerful. In other words the power given by the test in Section B serves as a lower bound for tests of H against K and the power given by the test in Section C serves as an upper bound for tests of H against K. It is seen that for large n and/or large alternative ρ the power of each of these tests is close to the power given by the simple test based on n_3. For example if n=100 and the alternative value of ρ is .05 then the power of the test based on n_3 alone which has α-level zero is .9941. If the marginal parameters were known and equal the power is .9948, .9954, .9962 for α-level tests of .05, .10, .20, respectively. There is not much gain in power in assuming the marginal parameters known and equal. As another example consider n=10 and the alternative value of ρ=.3. The power of the n_3 test is .9718 whereas if the marginal parameters were known and equal the power is .9784, .9816, .9860 for α-level tests of .05, .10, .20, respectively. The gain in power over the n_3 test is significant when n and ρ are small. For example, if n=20 and ρ=.1 the n_3 test gives a power of .878 whereas if the marginal parameters were known and equal the power is .894, .903, .920 for α-level tests of .05, .10, .20, respectively. If the marginal parameters were unknown but equal then the power is .886, .895, .914 for α-level tests of .05, .10, .20 respectively.

These results also indicate that one could not hope to improve much over the n_3 test with all parameters unknown even though it has α=0 since it is so close to the UMP test for the special case.

F. Confidence Regions for ρ and Marginal Parameters

Let $\{(X_j, Y_j)\}$ be a random sample of size n from BVE(γ_1, γ_2, ρ). Since n_3(number of (X_j, Y_j) such that $X_j=Y_j$) has a binomial distribution with parameters n and ρ, a confidence interval for ρ may be based on n_3. Tables are available to find numbers ρ_1 and ρ_2 as functions of n_3 and n such that (ρ_1, ρ_2) is a $100\alpha_1\%$ confidence interval for the parameter ρ. Also $2\lambda\sum\min(X_j,Y_j)$ has a Chi-Square distribution with 2n degrees of freedom as noted previously. So a confidence interval on λ may be obtained in terms of the statistic $AU=\sum\min(X_j,Y_j)$. For example, a 100α% two-sided confidence interval for λ is $(A/2AU, B/2AU)$ where $A=\chi_a^2(2n)$, $B=\chi_b^2(2n)$ and $a=(1-\alpha)/2$, $b=1-a$.

When $\rho=0$, $\lambda=\gamma_1+\gamma_2$. So a confidence interval on $\gamma_1+\gamma_2$ may be based on the statistic $AU=\sum\min(X_j,Y_j)$ when X and Y are uncorrelated. $(A/2AU, B/2AU)$ is a $100\alpha_2$% confidence interval for $\gamma_1+\gamma_2$ when $\rho=0$ where $A=\chi_a^2(2n)$, $B=\chi_b^2(2n)$ with $a=(1-\alpha_2)/2$ and $b=1-a$.

This same interval $(A/2AU, B/2AU)$ may be used as an approximate $100\alpha_2$% confidence interval for $\gamma_1+\gamma_2$ when $\rho>0$. However, since $E(AU) = n/\lambda=n(1+\rho)/(\gamma_1+\gamma_2)$ it is clear that the interval should be shifted to the right due to the factor $(1+\rho)$. Using this approximate confidence interval the actual confidence was computed by simulation. Table XIV compares the actual confidence with the approximate confidence for various ρ. Eight thousand samples of size 10, 4000 samples of size 20 and 2000 samples of size 40 were used.

TABLE XIV

ACTUAL CONFIDENCE FOR APPROXIMATE 100α% CONFIDENCE INTERVAL ON γ

γ	ρ	n	α=.95	α=.90	α=.80
1	.05	20	.940	.883	.781
		40	.947	.895	.793
		100	.949	.900	.795
1	.10	20	.937	.880	.768
		40	.945	.882	.773
		100	.948	.900	.790
		200	.952	.905	.798
1	.30	20	.899	.830	.720
		40	.908	.847	.732
		100	.928	.864	.755

If it is known that $\gamma_1=\gamma_2$ then an approximate $100\alpha_2\%$ confidence interval for γ, the common value of γ_1, γ_2, is (A/4AU, B/4AU).

Now $\lambda = (\gamma_1+\gamma_2)/(1+\rho)$ and $2\lambda AU \sim \chi^2(2n)$ together imply that (A/2AU, B/2AU) is really a $100\alpha_1\%$ confidence interval for $(\gamma_1+\gamma_2)/(1+\rho)$. Since AU and n_3 are independent, the region defined by $A/2AU < (\gamma_1+\gamma_2)/(1+\rho) < B/2AU$ and $\rho_1 < \rho < \rho_2$ is a joint $100\alpha_1\alpha_2\%$ confidence region for ρ and $\gamma_1+\gamma_2$. This will be illustrated for the case when $\gamma_1=\gamma_2=\gamma$. That is, the region defined by $A/4AU < \gamma/(1+\rho)<B/4AU$ and $\rho_1 < \rho < \rho_2$ is a joint $100\alpha_1\alpha_2\%$ confidence region for ρ and γ. Given n_3 and AU the confidence region appears as in Figure 2.

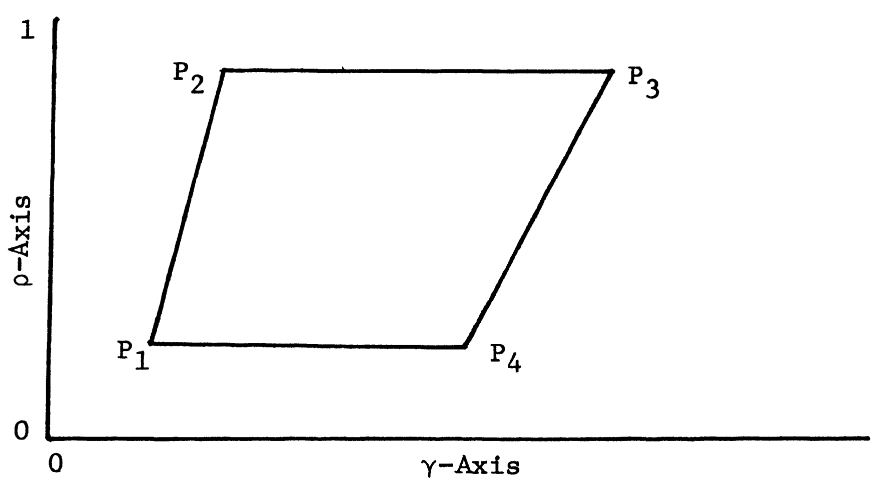

Figure 2. A Joint Confidence Region for ρ and γ.

The (γ, ρ) coordinates of the points P_1, P_2, P_3, P_4 are $P_1 = [A(1+\rho_1)/4AU, \rho_1]$, $P_2 = [A(1+\rho_2)/4AU, \rho_2]$, $P_3 = [B(1+\rho_2)/4AU, \rho_2]$, $P_4 = [B(1+\rho_1)/4AU, \rho_1]$. By using the γ values of points P_1 and P_3 it is clear that $[A(1+\rho_1)/4AU, B(1+\rho_2)/4AU]$ is at least a $100\alpha_1\alpha_2\%$ confidence interval for γ. By adjusting α_1 and α_2 such that $\alpha_1\alpha_2$ is fixed one may be able to shorten this confidence interval for γ. For the examples considered by the author it was found that there is not much gain over choosing $\alpha_1=\alpha_2$. To illustrate the method suppose n=20 and $n_3=1$ and AU=15. Take $\alpha_1=\alpha_2=.95$, so $\alpha_1\alpha_2=.9025$. Now ρ_1 and ρ_2 are found to be .01 and .25 so that (.01, .25) is a 95% confidence interval for ρ. A/4AU and B/4AU are respectively, .41 and .99 since $A = \chi^2_{.025}(40) = 24.4$ and $B = \chi^2_{.975}(40) = 59.3$. $A(1+\rho_1)/4AU = .42$ and $B(1+\rho_2)/4AU = 1.24$ so that (.42, 1.24) is at least a 90.25% confidence interval for γ. The true values of ρ and γ were .1 and 1, respectively. Shifting the approximate 95% confidence interval (.41, .99) for γ by $1+\rho = 1.1$ yields (.45, 1.09). In this example the approximate 95% confidence interval for γ did not cover γ. The procedure for one-sided and/or two-sided combinations for ρ and $\gamma_1+\gamma_2$ may be obtained in similar ways.

VII. RELIABILITY ESTIMATION

A. Introduction

Suppose X and Y are random variables with joint distribution function $F(x,y)$. If X and Y denote the life times of two components then it is natural to consider the reliability of a system with the two components in series or in parallel. In the case of two components in series, a success is defined by the event $X > c$ and $Y > c$ where c is a specified time. The series system is termed a success if both components survive at least time c. Hence the reliability of the two component series system is $P[X > c, Y > c] = P[\min(X,Y) > c] = \overline{F}(c,c)$. In the case of two components in parallel, a success is defined by the event $X > c$ or $Y > c$ where c is a specified time. The parallel system is termed a success if at least one component survives at least time c. Hence the reliability of the two component parallel system is $1 - P[X < c, Y < c] = P[\max(X,Y) > c] = 1 - F(c,c)$.

If X and Y are assumed independent when in fact they are dependent the estimation of these reliabilities may be quite misleading. Suppose $(X,Y) \sim BVE(\gamma_1, \gamma_2, \rho)$ with $\rho > 0$, but it is assumed that $(X,Y) \sim BVE(\gamma_1, \gamma_2, 0)$ or equivalently X and Y are independent and each has a marginal distribution which is an exponential with parameters γ_1, γ_2, respectively. In Section D examples are given to compare the reliability estimation based on the assumption of independence when in truth the random variables are distrubuted as $BVE(\gamma_1, \gamma_2, \rho)$ with $\rho > 0$. Section E gives another illustration (based on the concept of a tolerance region) of the discrepancy between an independent analysis when the random variables are positively correlated. In Section B a point estimate for the series reliability and an exact lower confidence limit for the series reliability is given. In Section C a point estimate for the parallel reliability and an asymptotic lower confidence limit for the parallel reliability is given.

B. Two Component Series System

Suppose $(X,Y) \sim BVE(\gamma_1, \gamma_2, \rho)$ and X denotes the lifetime of component I and Y denotes the lifetime of component II where the two components are in series. Then the reliability of this series system is $R_s = P[\min(X,Y) > c]$ where a system success is defined as the event $\min(X,Y) > c$. Now $\min(X,Y)$ has an exponential distribution with parameter $\lambda = (\gamma_1+\gamma_2)/(1+\rho)$ so $R_s = \exp\{-\lambda c\}$. A point estimate $\hat{R}_s$ of R_s is obtained by replacing the parameter λ with a point estimate. That is, $\hat{R}_s = \exp\{-\hat{\lambda} c\}$. There are at least three point estimates of R_s available since $\hat{\lambda}$ may be found by each of the three methods MLE, PROP or ACE referred to in Chapters IV and V. These will be compared in Section D.

Since R_s is a monotone function of λ, a lower confidence limit for R_s is available if an upper confidence limit can be found for λ. Such is the case since $2\lambda \sum \min(X_j, Y_j) = 2\lambda AU$ has a Chi-square distribution with 2n degrees of freedom. Hence $P[0 < 2\lambda AU < \chi^2_\alpha(2n)] = \alpha$ implies that $A = \chi^2_\alpha(2n)/2AU$ is an upper $100\alpha\%$ confidence limit for λ where AU is based on a random sample of size n from $BVE(\gamma_1, \gamma_2, \rho)$. Now $0 < \lambda < A$ implies $\exp\{-Ac\} < \exp\{-\lambda c\} = R_s < 1$ so that $\exp\{-Ac\}$ is a lower $100\alpha\%$ confidence limit for R_s.

It should be noted that X and Y need not represent lifetimes of the components but could represent any pair of responses which have a BVED and success is defined as $\min(X,Y) > c$. For example X and Y may represent two different responses from the same component. The analogous remark should be clear for the next section concerning the two component parallel system where success is defined by $\max(X,Y) > c$.

C. Two Component Parallel System

Suppose $(X,Y) \sim BVE(\gamma_1, \gamma_2, \rho)$ and X denotes the lifetime of component I and Y denotes the lifetime of component II where the two components are in parallel. Then the reliability of this two component parallel system is $R_p = P[\max(X,Y) > c]$ where a system success is defined as the event $\max(X,Y) > c$. $P[\max(X,Y) > c]$ was found in Chapter III to be the expression $\exp\{-\gamma_1 c\} + \exp\{-\gamma_2 c\} - \exp\{-\lambda c\}$ where $\lambda = (\gamma_1+\gamma_2)/(1+\rho)$. Hence a point estimate of R_p is $\hat{R}_p = \exp\{-\hat{\gamma}_1 c\} + \exp\{-\hat{\gamma}_2 c\} - \exp\{-\hat{\lambda} c\}$ where $\hat{\gamma}_1$, $\hat{\gamma}_2$, $\hat{\lambda}$ are point estimates of γ_1, γ_2, λ, respectively. Again at least three point estimates of R_p are available from the three estimation methods, MLE, PROP and ACE. These will be compared in Section D.

Obtaining a lower confidence limit for R_p presents a problem since R_p is not a monotone function of a single parameter. However, an asymptotic result may be obtained in the following manner. Let $R_p = g(\lambda_1, \lambda_2, \lambda_3)$ where $\lambda_1 = (\gamma_1-\rho\gamma_2)/(1+\rho)$, $\lambda_2 = (\gamma_2-\rho\gamma_1)/(1+\rho)$, $\lambda_3 = \rho(\gamma_1+\gamma_2)/(1+\rho)$ and $\lambda = (\gamma_1+\gamma_2)/(1+\rho)$. Then $\hat{R}_p = g(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3)$. Now $E(\hat{R}_p) = R_p + O(1/n)$ where the estimates are based on a sample of size n. That is, $\hat{R}_p$ is consistent. Expanding the function g through linear terms about the point $(\lambda_1, \lambda_2, \lambda_3)$ it is found that $\hat{R}_p = R_p + g_1(\hat{\lambda}_1-\lambda_1) + g_2(\hat{\lambda}_2-\lambda_2) + g_3(\hat{\lambda}_3-\lambda_3)$ + terms of degree two or greater in $(\hat{\lambda}_1-\lambda_1)$, $(\hat{\lambda}_2-\lambda_2)$, $(\hat{\lambda}_3-\lambda_3)$ where g_i denotes the partial of g with respect to its i^{th} argument evaluated at $(\lambda_1, \lambda_2, \lambda_3)$, i=1,2,3. By taking the expected value of $(\hat{R}_p-R_p)$ one obtains

$$E(\hat{R}_p-R_p)^2 = g_1^2 Var(\hat{\lambda}_1) + g_2^2 Var(\hat{\lambda}_2) + g_3^2 Var(\hat{\lambda}_3) + 2g_1g_2 Cov(\hat{\lambda}_1,\hat{\lambda}_2) +$$
$$2g_1g_3 Cov(\hat{\lambda}_1,\hat{\lambda}_3) + 2g_2g_3 Cov(\hat{\lambda}_2,\hat{\lambda}_3) + O(n^{-3/2}).$$

By replacing $Cov(\hat{\lambda}_i,\hat{\lambda}_j)$ with the corresponding term in the covariance matrix of the asymptotic distribution of $(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3)$ in the above expression one obtains $Var(\hat{R}_p)$ as a function of λ_1, λ_2, λ_3, alone, call it $h(\lambda_1, \lambda_2, \lambda_3)$, plus terms of order $n^{-3/2}$. So $n\cdot Var(\hat{R}_p)$ is asymptotically $n\cdot h(\lambda_1, \lambda_2, \lambda_3)$. An estimate of $Var(\hat{R}_p)$ is then $h(\hat{\lambda}_1,\hat{\lambda}_2,\hat{\lambda}_3)$. An approximate $100(1-\alpha)\%$ lower confidence limit for R_p then is $\hat{R}_p - z_\alpha[h(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3)]^{1/2}$ where z_α is the normal abscissa such that the integral of the standard normal density from z_α to ∞ is α. In Table XV below the actual confidences when $R_p = .75, .90, .95$ are compared to the desired level based on simulations for n=20, 40, 100. As n increases the actual confidence approaches the desired level.

D. Effect of Erroneous Assumption of Independence

Suppose $(X,Y) \sim BVE(\gamma_1, \gamma_2, \rho)$ where $\rho > 0$. On the basis of a random sample of size n from this distribution a point estimate and a lower $100\alpha\%$ confidence limit is required for R_s or R_p. If the assumption is made that X and Y are independent random variables each with an exponential distribution with parameters γ_1, γ_2, respectively, the estimation of R_s or R_p would be in error. In general R_s will be underestimated and R_p will be overestimated. Since $R_s = \exp\{-\lambda c\}$ where $\lambda = (\gamma_1+\gamma_2)/(1+\rho)$, by assuming erroneously that $\rho=0$ the estimate for R_s would be based on an estimate for just $\gamma_1 + \gamma_2$ which is in general larger than that for $(\gamma_1+\gamma_2)/(1+\rho)$. Hence the independent analysis would lead to an underestimate of R_s with a corresponding conservative lower confidence limit for R_s. From the form of

TABLE XV

ACTUAL CONFIDENCE FOR APPROXIMATE 100α% LOWER CONFIDENCE LIMIT ON R_p

				R_p=.75		R_p=.90		R_p=.95	
γ	ε	ρ	n	α=.80	α=.95	α=.80	α=.95	α=.80	α=.95
1.0	1.0	.05	20	810	938	797	912	751	860
			40	826	946	802	936	786	906
			100	819	940	809	936	805	924
1.0	1.0	.10	20	818	935	790	900	765	869
			40	818	939	796	921	779	912
			100	830	953	806	938	786	927
1.0	.5	.10	40	886	982	846	952	816	930
			100	885	990	833	965	810	946
1.0	1.0	.30	20	820	941	803	932	798	929
			40	809	945	800	936	799	935
			100	818	945	812	944	800	939

$R_p = \exp\{-\gamma_1 c\} + \exp\{-\gamma_2 c\} - \exp\{-\lambda c\}$, the erroneous assumption of $\rho=0$ would lead to an overestimate of R_p with a corresponding over-optimistic lower confidence limit for R_p.

When $\rho=0$ is assumed then the distribution of (X,Y) is considered to have density $f(x,y) = \gamma_1\gamma_2\exp\{-\gamma_1 x-\gamma_2 y\}$. The maximum likelihood estimates for γ_1 and γ_2 based on a sample of size n are $\hat{\gamma}_1 = n/\sum X_j$ and $\hat{\gamma}_2 = n/\sum Y_j$. When $\rho=0$ is assumed, $\hat{R}_s$ is computed by $\exp\{-(\hat{\gamma}_1+\hat{\gamma}_2)c\}$ with $\hat{\gamma}_1$ and $\hat{\gamma}_2$ the maximum likelihood estimates given above. Similarly when $\rho=0$ is assumed, $\hat{R}_p$ is computed by $\exp\{-\hat{\gamma}_1 c\} + \exp\{-\hat{\gamma}_2 c\} - \exp\{-(\hat{\gamma}_1+\hat{\gamma}_2)c\}$.

The number c was chosen to give an expected reliability of .75, .90 or .95. This was accomplished by solving for c in the equation $R_s = \exp\{-\lambda c\}$ or $R_p = \exp\{-\gamma_1 c\} + \exp\{-\gamma_2 c\} - \exp\{-\lambda c\}$ where $\lambda = (\gamma_1+\gamma_2)/(1+\rho)$. Then samples of size n were generated from BVE(γ_1,γ_2,ρ) and $\hat{R}_s$ and $\hat{R}_p$ were computed for each sample by the methods of MLE, PROP, ACE not assuming $\rho=0$ and also by the method in the preceding paragraph where $\rho=0$ was assumed, that is, the independent analysis. Also 80% and 95% lower confidence limits for R_s and R_p were computed for each of the four methods from the simulated distributions of $\hat{R}_s$ and $\hat{R}_p$.

Tables XVI and XVII compare the simulated mean of $\hat{R}_s$ and $\hat{R}_p$ for the three estimation methods MLE, PROP and ACE with the mean of $\hat{R}_s$ and $\hat{R}_p$ based on the assumption of independence. The numbers are based on simulations. For n=10, 8000 samples were used; for n=20, 4000 samples were used; for n=40, 2000 samples were used; for n=100, 1000 samples were used.

TABLE XVI

SERIES RELIABILITY ESTIMATES

γ	ε	ρ	n	Dependent Analysis MLE	PROP	ACE	Independent MLE	True R
1.0	1.0	.10	10	.726	.747	.751	.704	.75
				.888	.897	.899	.878	.90
				.943	.948	.949	.938	.95
			20	.740	.750	.751	.718	.75
				.895	.900	.900	.885	.90
				.947	.950	.950	.942	.95
			40	.743	.749	.749	.722	.75
				.897	.899	.899	.887	.90
				.948	.950	.950	.943	.95
			100	.747	.749	.749	.726	.75
				.898	.899	.899	.889	.90
				.949	.950	.950	.944	.95
1.0	1.0	.30	10	.727	.748	.752	.664	.75
				.888	.898	.899	.859	.90
				.943	.948	.949	.928	.95
			20	.740	.750	.752	.677	.75
				.895	.900	.900	.866	.90
				.947	.950	.950	.932	.95
			40	.744	.749	.750	.681	.75
				.897	.899	.900	.869	.90
				.948	.950	.950	.934	.95
			100	.747	.749	.749	.685	.75
				.899	.900	.900	.871	.90
				.949	.950	.950	.935	.95
1.0	1.0	.05	10	.727	.748	.751	.716	.75
				.888	.898	.899	.884	.90
				.943	.948	.949	.941	.95
			20	.740	.751	.752	.729	.75
				.895	.900	.900	.890	.90
				.947	.950	.950	.945	.95
			40	.743	.749	.748	.732	.75
				.897	.899	.899	.892	.90
				.948	.950	.949	.946	.95
			100	.748	.750	.750	.737	.75
				.900	.900	.900	.894	.90
				.950	.950	.950	.947	.95

TABLE XVI Continued

γ	ε	ρ	n	Dependent Analysis MLE	PROP	ACE	Independent MLE	True R
1.0	.5	.10	10	.728	.749	.754	.707	.75
				.889	.898	.900	.879	.90
				.944	.949	.949	.939	.95
			20	.738	.749	.751	.717	.75
				.894	.899	.900	.885	.90
				.947	.949	.950	.942	.95
			40	.744	.750	.751	.723	.75
				.897	.900	.900	.888	.90
				.949	.950	.950	.944	.95
			100	.748	.750	.751	.727	.75
				.899	.900	.900	.890	.90
				.950	.950	.950	.945	.95

TABLE XVII

PARALLEL RELIABILITY ESTIMATES

γ	ε	ρ	n	Dependent Analysis MLE	Dependent Analysis PROP	Dependent Analysis ACE	Independent MLE	True R_p
1.0	1.0	.10	10	.726	.756	.765	.759	.75
				.886	.900	.902	.917	.90
				.942	.949	.949	.966	.95
			20	.740	.756	.758	.773	.75
				.895	.902	.902	.926	.90
				.947	.951	.950	.971	.95
			40	.743	.751	.752	.777	.75
				.896	.900	.900	.927	.90
				.948	.950	.950	.972	.95
			100	.748	.751	.750	.781	.75
				.899	.900	.900	.929	.90
				.950	.950	.950	.972	.95
1.0	1.0	.30	10	.730	.757	.760	.823	.75
				.889	.901	.901	.958	.90
				.944	.950	.950	.987	.95
			20	.741	.755	.756	.836	.75
				.896	.901	.901	.962	.90
				.948	.951	.950	.989	.95
			40	.744	.751	.752	.839	.75
				.897	.900	.900	.964	.90
				.948	.950	.950	.989	.95
			100	.748	.751	.751	.843	.75
				.899	.900	.900	.965	.90
				.950	.950	.950	.990	.95
1.0	0.5	0.1	10	.729	.759	.768	.762	.75
				.888	.902	.903	.920	.90
				.943	.950	.950	.968	.95
			20	.737	.753	.758	.771	.75
				.893	.901	.901	.925	.90
				.946	.950	.950	.971	.95
			40	.744	.752	.755	.778	.75
				.897	.901	.901	.929	.90
				.948	.950	.950	.973	.95
			100	.748	.751	.752	.782	.75
				.899	.900	.900	.931	.90
				.949	.950	.950	.973	.95

Tables XVIII and XIX compare the lower 80% and 95% confidence limits for R_s and R_p for each of the methods MLE, PROP, ACE with that of the independent analysis.

By examining these tables the following observations may be made. For the dependent analysis the method of estimation using MLE or PROP is superior to that of ACE since the variance of the MLE and PROP estimates for R_s and R_p is smaller than that of the ACE estimates for R_s and R_p. Consequently the lower confidence limits for R_s and R_p are tighter for MLE and PROP estimation than for ACE estimation.

The independent analysis underestimates R_s and overestimates R_p. The discrepancy of the under or over estimation becomes greater as the sample size n increases or as the value of ρ increases. The discrepancy becomes so severe that in the example of $(\gamma_1, \gamma_2, \rho) = (1, 1, .1)$ with sample size n=40 the independent analysis leads to a 95% lower confidence limit for R_p=.90 of .900 and an 80% lower confidence limit for R_p=.90 of .915 which are obviously over-optimistic. That is, the true value for R_p is .90 and yet a sample of size 40 will consistently give 80% and 95% lower confidence limits which exceed the true value. This illustrates that using an independent analysis when in fact $\rho>0$ may be very misleading especially in the parallel situation. In the series situation the independent analysis is too conservative in the sense that the correct dependent analysis would lead to more information. For example, consider again $(\gamma_1, \gamma_2, \rho) = (1, 1, .1)$ and n=40. A lower 95% confidence limit for R_s=.75 is expected to be .703 using PROP whereas the independent analysis only yields on the average .672.

TABLE XVIII

COMPARISON OF 100α% LOWER CONFIDENCE LIMITS FOR R_s

γ	ε	ρ	n	Dependent Analysis MLE	PROP	ACE	Independent MLE	True R_s	α
1.0	1.0	0.1	20	.676	.688	.664	.647	.75	.95
				.709	.721	.712	.686		.80
				.865	.871	.860	.852	.90	.95
				.882	.886	.883	.871		.80
				.932	.933	.929	.924	.95	.95
				.941	.942	.941	.934		.80
			40	.697	.703	.686	.672	.75	.95
				.722	.727	.720	.699		.80
				.875	.879	.871	.863	.90	.95
				.886	.890	.887	.876		.80
				.935	.938	.933	.931	.95	.95
				.942	.943	.942	.937		.80
			100	.721	.722	.712	.697	.75	.95
				.734	.736	.732	.712		.80
				.884	.886	.881	.874	.90	.95
				.892	.893	.891	.882		.80
				.941	.941	.940	.934	.95	.95
				.943	.944	.944	.941		.80
1.0	1.0	0.3	20	.670	.682	.662	.600	.75	.95
				.709	.720	.713	.638		.80
				.863	.869	.860	.823	.90	.95
				.881	.886	.883	.848		.80
				.931	.932	.929	.910	.95	.95
				.940	.942	.941	.922		.80
			40	.694	.700	.684	.618	.75	.95
				.721	.726	.721	.653		.80
				.874	.877	.870	.838	.90	.95
				.886	.890	.886	.855		.80
				.935	.937	.933	.917	.95	.95
				.942	.943	.942	.925		.80
			100	.718	.720	.711	.651	.75	.95
				.733	.735	.732	.669		.80
				.883	.885	.881	.852	.90	.95
				.892	.893	.892	.863		.80
				.941	.941	.940	.923	.95	.95
				.943	.944	.944	.931		.80

TABLE XVIII Continued

γ	ε	ρ	n	Dependent Analysis MLE	Dependent Analysis PROP	Dependent Analysis ACE	Independent MLE	True R_s	α
1.0	1.0	.05	20	.674	.687	.664	.661	.75	.95
				.710	.722	.713	.697		.80
				.865	.871	.861	.860	.90	.95
				.882	.887	.883	.876		.80
				.931	.933	.930	.928	.95	.95
				.941	.942	.941	.937		.80
			40	.697	.704	.685	.685	.75	.95
				.722	.728	.720	.710		.80
				.875	.879	.871	.870	.90	.95
				.888	.890	.886	.882		.80
				.935	.938	.933	.932	.95	.95
				.942	.943	.942	.941		.80
			100	.720	.722	.713	.708	.75	.95
				.734	.736	.732	.723		.80
				.884	.885	.882	.881	.90	.95
				.892	.893	.892	.887		.80
				.941	.941	.941	.940	.95	.95
				.944	.944	.944	.942		.80
1.0	0.5	0.1	20	.667	.680	.661	.639	.75	.95
				.706	.718	.712	.682		.80
				.862	.867	.859	.848	.90	.95
				.880	.885	.883	.869		.80
				.930	.932	.928	.922	.95	.95
				.940	.942	.941	.933		.80
			40	.695	.701	.690	.672	.75	.95
				.722	.728	.724	.699		.80
				.874	.878	.872	.863	.90	.95
				.887	.890	.889	.877		.80
				.935	.938	.934	.931	.95	.95
				.942	.943	.943	.937		.80
			100	.721	.722	.714	.696	.75	.95
				.734	.736	.732	.712		.80
				.884	.885	.882	.874	.90	.95
				.892	.893	.892	.882		.80
				.941	.941	.941	.934	.95	.95
				.944	.944	.944	.941		.80

TABLE XIX

COMPARISON OF 100α% LOWER CONFIDENCE LIMIT ON R_p

γ	ε	ρ	n	Dependent Analysis MLE	PROP	ACE	Independent MLE	True R_p	α
1.0	1.0	0.1	20	.632	.651	.612	.678	.75	.95
				.693	.709	.695	.731		.80
				.835	.844	.824	.886	.90	.95
				.868	.876	.870	.910		.80
				.911	.915	.904	.953	.95	.95
				.931	.935	.932	.963		.80
			40	.666	.675	.646	.708	.75	.95
				.707	.716	.705	.745		.80
				.854	.859	.844	.900	.90	.95
				.876	.881	.876	.915		.80
				.921	.923	.917	.960	.95	.95
				.935	.938	.936	.965		.80
			100	.704	.706	.688	.743	.75	.95
				.727	.731	.723	.763		.80
				.875	.876	.868	.913	.90	.95
				.887	.888	.886	.922		.80
				.933	.934	.931	.962	.95	.95
				.942	.943	.941	.969		.80
1.0	1.0	0.3	20	.637	.649	.618	.750	.75	.95
				.694	.707	.699	.800		.80
				.842	.847	.833	.938	.90	.95
				.872	.878	.874	.953		.80
				.917	.921	.913	.980	.95	.95
				.934	.938	.935	.983		.80
			40	.666	.671	.653	.777	.75	.95
				.710	.716	.708	.824		.80
				.858	.861	.852	.946	.90	.95
				.880	.883	.879	.956		.80
				.927	.928	.924	.981	.95	.95
				.939	.940	.939	.984		.80
			100	.705	.709	.696	.811	.75	.95
				.728	.730	.724	.886		.80
				.876	.879	.873	.953	.90	.95
				.889	.890	.887	.961		.80
				.936	.936	.934	.981	.95	.95
				.943	.943	.943	.984		.80

TABLE XIX Continued

γ	ε	ρ	n	Dependent Analysis MLE	PROP	ACE	Independent MLE	True R_p	α
1.0	0.5	0.1	20	.627	.647	.601	.672	.75	.95
				.690	.705	.688	.728		.80
				.831	.840	.821	.885	.90	.95
				.866	.874	.867	.910		.80
				.907	.913	.903	.953	.95	.95
				.930	.934	.931	.963		.80
			40	.672	.682	.653	.713	.75	.95
				.710	.719	.707	.748		.80
				.857	.861	.847	.903	.90	.95
				.879	.883	.877	.917		.80
				.923	.924	.919	.961	.95	.95
				.937	.940	.936	.966		.80
			100	.708	.711	.688	.743	.75	.95
				.727	.731	.722	.765		.80
				.875	.876	.866	.915	.90	.95
				.887	.888	.885	.923		.80
				.933	.933	.931	.964	.95	.95
				.942	.942	.942	.971		.80

E. An Illustration Based on a Tolerance Region Approach

Consider a two component system and suppose $(X,Y)\sim BVE(\gamma_1, \gamma_2, \rho)$ with $\rho>0$ where X denotes a response from one component and Y denotes a response from the second component. Suppose further that a system success is defined by the event $S=\{X\geq a_1 \text{ and } Y\geq a_2\}$ where a_1 and a_2 are specified. S will be referred to as the specification region. For example, if X and Y denote life times of the components, a system success is such that the first component survives at least time a_1 and the second survives at least a_2. For the special case when $a_1=a_2$ S could be interpreted as success for a two component series system. Then the reliability R for the system is the probability that $(X,Y)\in S$. Now

$$R = P[(X,Y) \in S] = \iint_S dF(x,y) = \overline{F}(a_1,a_2) = \exp\{-\lambda_1 a_1-\lambda_2 a_2-\lambda_3 \max(a_1,a_2)\}.$$

As in the previous sections an estimate of R may be obtained by replacing the unknown parameters in the above expression with estimates. However a different approach is taken here.

In [16] Bain considers the evaluation of the reliability of a system of components where the components follow a joint exponential distribution through the approach of a tolerance region. A set T is called an (α,R) tolerance region if $P\{P[(X,Y) \in T] \geq R\} = \alpha$. This means that with α confidence at least 100R% of the population falls in T. If such an (α,R) tolerance region T is calculated on the basis of a sample and T is contained in the specification set S, then with at least 100α% confidence the reliability of the system is at least R. In the case where X and Y are considered independent each with exponential marginals with parameters γ_1 and γ_2, respectively then such an (α,R) tolerance region T has been developed in [16]. T has the form $T = \{(X,Y) | X \geq cn\overline{X},\ Y \geq cn\overline{Y}$ where $c = -2\ \ln(R)/\chi^2_\alpha(4n)$. The method to be considered in this section will be as follows. Given α and R from the above independent analysis the actual confidence α' will be found for the same R by simulating the $P\{P[(X,Y) \in T] \geq R\}$ where T is chosen as if X and Y were independent but the samples will come from $BVE(\gamma_1,\ \gamma_2,\ \rho)$ with $\rho>0$. In this way a comparison of the confidence α given by the independent analysis will be made with the actual confidence α' when ρ is in fact positive. It will be noted that α' is not a function of R. Also a comparison will be made by keeping the confidence α fixed and computing the actual content R' when the samples come from $BVE(\gamma_1,\ \gamma_2,\ \rho)$ with $\rho>0$. Here it appears that R' is independent of n and α. These comparisons are made in Tables XX and XXI. The following observation should be made. The independent analysis gives conservative values for R and α compared to the corresponding dependent analysis. That is, if in fact $\rho>0$ but yet the assumption is independence ($\rho=0$), then the values for α and R are smaller than they actually should be.

TABLE XX

DEPENDENT α FOR FIXED R COMPARED TO INDEPENDENT α

			Independent		Dependent α		
γ	ε	ρ	α	n=10	n=20	n=40	n=100
1.0	1.0	.05	.80	.848	.867	.896	.944
			.95	.968	.971	.981	.987
1.0	0.5	.10	.80	.869	.897	.933	.970
			.95	.970	.978	.991	.994
1.0	1.0	.10	.80	.878	.905	.940	.984
			.95	.974	.983	.992	.998
1.0	1.0	.30	.80	.956	.983	.997	1.000
			.95	.992	.998	1.000	1.000

TABLE XXI

DEPENDENT R FOR FIXED α COMPARED TO INDEPENDENT R

γ	ε	ρ	Independent R	Dependent R n=10	n=20	n=40	n=100
1.0	1.0	.05	.75	.758	.759	.760	.761
			.90	.903	.904	.905	.905
			.95	.952	.952	.952	.952
1.0	0.5	0.1	.75	.762	.764	.764	.765
			.90	.905	.906	.906	.907
			.95	.952	.953	.953	.953
1.0	1.0	0.1	.75	.765	.766	.767	.768
			.90	.906	.907	.907	.907
			.95	.953	.954	.954	.954
1.0	1.0	0.3	.75	.788	.790	.794	.798
			.90	.917	.918	.919	.920
			.95	.958	.959	.959	.960

The details of the manner in which these numbers α' and R' given in Tables XX and XXI are obtained are given now. Consider first the numbers α'. Set α and R and consider T to be the tolerance region given in [16] of the form $T = \{(X,Y)|X \geq c n\overline{X},\ Y \geq c n\overline{Y}\}$ where $c = 2\ \ln(R)/\chi^2_\alpha(4n)$ and $\overline{X}$, $\overline{Y}$ are the means based on a sample of size n from BVE(γ_1, γ_2, ρ) with $\rho>0$ or equivalently BVE(λ_1, λ_2, λ_3) with $\lambda_3>0$. α' is computed by $\alpha' = P\{P[X,Y) \in T] \geq R\}$ or equivalently $\alpha' = P[\lambda_1\overline{X}+\lambda_2\overline{Y}+\lambda_3\max(\overline{X},\overline{Y}) \leq -\ln(R)/cn = \chi^2_\alpha(4n)/2n]$. From the above expression it is clear that α' is not a function of R but only dependent on α and n along with the parameters of the distribution. The probability α' was then found by generating many samples of size n from BVE(λ_1, λ_2, λ_3) and counting the proportion of times that $\lambda_1\overline{X}$ + $\lambda_2 Y$ + $\lambda_3\max(\overline{X},\overline{Y})$ was less than $\chi^2_\alpha(4n)/2n$.

Now consider the numbers R'. Set α and R and consider T to be the tolerance region given above. For the given α the number h is found such that $P[\lambda_1\overline{X} + \lambda_2\overline{Y} + \lambda_3\max(\overline{X},\overline{Y}) \leq h] = \alpha$. The distribution of $\lambda_1\overline{X} + \lambda_2\overline{Y} + \lambda_3\max(\overline{X},\overline{Y})$ is simulated where $\overline{X}$ and $\overline{Y}$ are based on a sample of size n from BVE(λ_1, λ_2, λ_3) with $\lambda_3>0$ (or equivalently $\rho>0$) and then h is the abscissa such that 100α% of the distribution lies to the left of h. Then h is set equal to $-\ln(R')/cn$ where $c = -2\ln(R)/\chi^2_\alpha(4n)$ from which R' is computed to be $R' = \exp\{2nh\ln(R)/\chi^2_\alpha(4n)\}$.

VIII. SUMMARY, CONCLUSIONS AND FURTHER PROBLEMS

The method of maximum likelihood estimation for the parameters of the bivariate exponential distribution is asymptotically efficient. The estimates are obtainable by solving a nonlinear system which is an easy matter with the aid of a computer. The PROP method of estimation is almost as efficient as MLE with the advantage that the estimates are computationally easier to obtain. Both MLE and PROP are superior to ACE, the method given by Arnold in [2], in terms of efficiency, even though ACE is unbiased and the other two methods are not. The estimation of the correlation ρ between X and Y based on the proportion of sample points such that X=Y is superior to the usual sample correlation coefficient.

Procedures have been developed to test whether $\rho=0$ or $\rho>0$. These tests are felt to be important since if $(X,Y) \sim BVE(\gamma_1, \gamma_2, \rho)$ errors are introduced when assuming wrongly that the random variables X and Y are independent ($\rho=0$) when in fact they are dependent ($\rho>0$). These errors are illustrated in the chapter on Reliability Estimation.

The tests developed for independence against dependence are interesting in themselves in that the power of the n_3 test, which has probability of type I error zero, serves as a uniform lower bound for the other tests considered and yet the power of the other tests approach that of the n_3 test as n increases or as ρ increases. One may say that the n_3 test is uniformly most powerful asymptotically [17].

The joint confidence region approach for ρ and γ when the marginal parameters are assumed equal gives a conservative confidence interval for γ. The exploitation of the fact that $2\lambda\sum \min(X_j,Y_j)$ has a chi-square distribution with 2n degrees of freedom led to confidence limits for λ and the series reliability $\exp\{-\lambda c\}$ as well as aiding in the computation of the power and the critical values for the tests considered in Chapter VI.

There is the question of whether the bivariate exponential distribution is a member of the exponential family of distributions or not. The matter of the existence of a minimal set of sufficient statistics for the parameters of the bivariate exponential distribution has not been considered.

The matter of extending the statistical procedures mentioned here to the multivariate exponential distribution including reliability of n component systems with dependencies built in needs further investigation.

Also the manner in which one handles the procedures of maximum likelihood estimation and hypothesis testing for the general class of distributions which are neither absolutely continuous nor discrete needs further investigation.

APPENDIX A

DERIVATION OF THE MAXIMUM LIKELIHOOD EQUATIONS

$(X,Y) \sim BVE(\lambda_1, \lambda_2, \lambda_3)$ if

$$\overline{F}(x,y) = P[X>x, Y>y] = \exp\{-\lambda_1 x - \lambda_2 y - \lambda_3 \max(x,y)\}$$

$$\lambda_1 > 0,\ \lambda_2 > 0,\ \lambda_3 \geq 0,\ x \geq 0,\ y \geq 0.$$

Following the notation in Wilks [10] Chapter 12, let

$$S_i(x,y;\ \lambda_1, \lambda_2, \lambda_3) =$$

$$\lim_{\substack{h\to o^+ \\ k\to o^+}} \left(\frac{\partial}{\partial\lambda_i}[dF(x,y;\ \lambda_1, \lambda_2, \lambda_3)]/dF(x,y;\ \lambda_1, \lambda_2, \lambda_3)\right) \qquad (1)$$

$(i=1,2,3)$ where $dF(x,y;\ \lambda_1, \lambda_2, \lambda_3) = F(x,y;\ \lambda_1, \lambda_2, \lambda_3) +$

$$F(x+h,\ y+k;\ \lambda_1, \lambda_2, \lambda_3) - F(x,y+k;\ \lambda_1, \lambda_2, \lambda_3) - F(x+h,y;\lambda_1, \lambda_2, \lambda_3).$$

Since $dF = d\overline{F}$ all F's above may be replaced by $\overline{F}$'s. Now $d\overline{F}(x,x) =$ $\exp\{-\lambda x\}[\exp\{-\lambda_1 h-\lambda_2 k-\lambda_3 \max(h,k)\} + 1 - \exp\{-\lambda_1 h-\lambda_3 h\} - \exp\{-\lambda_2 k-\lambda_3 k\}]$.

In the regions where the distribution F possesses a density f the expression for S_i may be simplified to

$$S_i(x,y;\ \lambda_1, \lambda_2, \lambda_3) = \frac{\partial}{\partial\lambda_i}\log f(x,y;\ \lambda_1, \lambda_2, \lambda_3) \qquad (i=1,2,3). \quad (2)$$

In the first quadrant let R1 denote the region where $x < y$, R2 denote the region where $x > y$ and R3 denote the ray where $x=y$. In the regions R1 and R2 the definition (2) will be utilized whereas in R3 the definition (1) will be used to obtain the expressions for S_i, $i=1,2,3$.

In R1 $f(x,y;\ \lambda_1,\ \lambda_2,\ \lambda_3) = \log(\lambda_1) + \log(\lambda_2+\lambda_3) - \lambda_1 x - \lambda_2 y - \lambda_3 y$ and in R2 $f(x,y;\ \lambda_1,\ \lambda_2,\ \lambda_3) = \log(\lambda_2) + \log(\lambda_1+\lambda_3) - \lambda_1 x - \lambda_2 y - \lambda_3 x$.

Hence

$$S_1(x,y) = \begin{cases} 1/\lambda_1 - x & (x,y)\varepsilon R1 \\ 1/(\lambda_1+\lambda_3)-x & (x,y)\varepsilon R2 \end{cases}$$

$$S_2(x,y) = \begin{cases} 1/(\lambda_2+\lambda_3) - y & (x,y)\varepsilon R1 \\ 1/\lambda_2 - y & (x,y)\varepsilon R2 \end{cases}$$

$$S_3(x,y) = \begin{cases} 1/(\lambda_2+\lambda_3) - y & (x,y)\varepsilon R1 \\ 1/(\lambda_1+\lambda_3) - x & (x,y)\varepsilon R2 \end{cases}$$

Let $A = \exp\{-\lambda_1 h-\lambda_3 h\}$, $B = \exp\{-\lambda_1 h-\lambda_2 k-\lambda_3 \max(h,k)\}$ and $C=\exp\{-\lambda_2 k-\lambda_3 k\}$.

So $dF(x,x) = \exp\{-\lambda x\}(1+B-A-C)$

$$\frac{\partial}{\partial\lambda_1}[dF(x,x)] = \exp\{-\lambda x\}[-x(1 + B - A - C) - h(B - A)]$$

$$\frac{\partial}{\partial\lambda_1}[dF(x,x)]/dF(x,x) = -x - h(B - A)/(1 + B - A - C)$$

Expanding A, B and C in a MacLaurin Series about $h = 0$, $k = 0$ and letting D_m denote terms of degree m or higher in h and k we have

$$A = 1 - \lambda_1 h - \lambda_3 h + D_2$$

$$B = 1 - \lambda_1 h - \lambda_2 k - \lambda_3 \max(h,k) + D_2$$

$$C = 1 - \lambda_2 k - \lambda_3 k + D_2$$

$$1 + B - A - C = \lambda_3[h + k - \max(h,k)] + D_2$$

$$h(B-A) = \lambda_2 hk - \lambda_3 h + \lambda_3 h \max(h,k) + D_3$$

Now $h(B-A)/(1+B-A-C) = \begin{cases} (\lambda_2 hk-\lambda_3 h^2+\lambda_3 hk+D_3)/(\lambda_3 h+D_2) & \text{if } h \le k \\ (\lambda_2 hk-\lambda_3 h^2+\lambda_3 h^2+D_3)/(\lambda_3 k+D_2) & \text{if } h > k \end{cases}$

So clearly $\lim\limits_{\substack{h\to o^+ \\ k\to o^+}} \dfrac{h(B - A)}{1 + B - A - C} = 0$. Hence $S_1(x,x) = -x$. By symmetry it follows that $S_2(x,x) = -x$.

$$\frac{\partial}{\partial\lambda_3}[dF(x,x)] = \exp\{-\lambda x\}[-x(1 + B - A - C) - \max(h,k)B + hA + kC]$$

$$\frac{\partial}{\partial\lambda_3}[dF(x,x)]/dF(x,x) = -x + [hA + kC - \max(h,k)B]/(1 + B - A - C)$$

$$[hA-\max(h,k)B+C]/(1+B-A-C) = \begin{cases} (h+\lambda_1 hk-\lambda_1 h^2-\lambda_3 h^2+D_3)/(\lambda_3 h+D_2) & \text{if } h \le k \\ (k+\lambda_2 hk-\lambda_2 k^2-\lambda_3 k^2+D_3)/(\lambda_3 k+D_2) & \text{if } h > k \end{cases}$$

So $\lim\limits_{\substack{h\to o^+ \\ k\to o^+}} \left([hA - \max(h,k)B + C]/(1 + B - A - C)\right) = 1/\lambda_3$.

Hence $S_3(x,x) = -x + 1/\lambda_3$.

Summarizing we have

$$S_1(x,y) = \begin{cases} 1/\lambda_1 - x & (x,y)\varepsilon R1 \\ 1/(\lambda_1+\lambda_3) - x & (x,y)\varepsilon R2 \\ -x & (x,y)\varepsilon R3 \end{cases}$$

$$S_2(x,y) = \begin{cases} 1/(\lambda_2+\lambda_3) - y & (x,y)\varepsilon R1 \\ 1/\lambda_2 - y & (x,y)\varepsilon R2 \\ -x & (x,y)\varepsilon R3 \end{cases}$$

$$S_3(x,y) = \begin{cases} 1/(\lambda_2+\lambda_3) - y & (x,y)\varepsilon R1 \\ 1/(\lambda_1+\lambda_3) - x & (x,y)\varepsilon R2 \\ 1/\lambda_3 - x & (x,y)\varepsilon R3 \end{cases}$$

Let $\{(X_j,Y_j)\}$ be a random sample of size n from BVE(λ_1, λ_2, λ_3). Then the maximum likelihood equations for $\hat{\lambda}_1$, $\hat{\lambda}_2$, $\hat{\lambda}_3$ are obtained from

$$\sum_{j=1}^{n} S_i(X_j,Y_j;\ \hat{\lambda}_1,\ \hat{\lambda}_2,\ \hat{\lambda}_3) = 0 \qquad (i=1,2,3)$$

The hats on the λ_i have been suppressed for notational convenience below.

$$\sum S_1(X_j,Y_j) = \sum_{R1} (1/\lambda_1 - X_j) + \sum_{R2} (1/(\lambda_1+\lambda_3) - X_j) + \sum_{R3} - X_j = 0$$

$$= n_1/\lambda_1 - \sum_{R1} X_j + n_2/(\lambda_1+\lambda_3) - \sum_{R2} X_j - \sum_{R3} X_j = 0$$

$$\text{or} \qquad n_1/\lambda_1 + n_2/(\lambda_1+\lambda_3) = \sum X_j$$

Similarly $\sum S_2(X_j,Y_j) = \sum_{R1} (1/(\lambda_2+\lambda_3) - Y_j) + \sum_{R2} (1/\lambda_2 - Y_j) + \sum_{R3} - Y_j = 0$

$$\text{or } n_1/(\lambda_2+\lambda_3) + n_2/\lambda_2 = \sum Y_j$$

And $\sum S_3(X_j,Y_j) = \sum_{R1} (1/(\lambda_2+\lambda_3) - Y_j) + \sum_{R2} (1/(\lambda_1+\lambda_3) - X_j) + \sum_{R3} (1/\lambda_3 - X_j) = 0$

$$\text{or} \qquad n_1/(\lambda_2+\lambda_3) + n_2/(\lambda_1+\lambda_3) + n_3/\lambda_3 = \sum \max(X_j,Y_j).$$

Here n_i counts the number of sample points (X_j,Y_j) which fall in Ri, $i=1,2,3$.

APPENDIX B

VERIFICATION OF THE REGULARITY CONDITIONS

The bivariate distribution function $F(x, y; \lambda_1, \lambda_2, \lambda_3)$ is said to be regular with respect to its first partial λ_i-derivatives in $\Omega_3=(0, \infty) \times (0, \infty) \times (0, \infty)$ if $E[S_i(X, Y; \lambda_1, \lambda_2, \lambda_3)] = 0$ for $i = 1, 2, 3$. $F(x,y; \lambda_1, \lambda_2, \lambda_3)$ is said to be regular with respect to its second partial λ_i-derivatives in Ω_3 if $E[S_i(X, Y; \lambda_1, \lambda_2, \lambda_3) \cdot S_j(X, Y; \lambda_1, \lambda_2, \lambda_3)] + E[S_{ij}(X, Y; \lambda_1, \lambda_2, \lambda_3)] = 0$ for all $i,j = 1, 2, 3$ where $S_i(x, y; \lambda_1, \lambda_2, \lambda_3)$ is given in Appendix A and $S_{ij}(x, y; \lambda_1, \lambda_2, \lambda_3)$ denotes the partial derivative of $S_j(x, y; \lambda_1, \lambda_2, \lambda_3)$ with respect to λ_i. The object of this appendix is to verify that the bivariate exponential distribution satisfies both of these regularity conditions.

The following facts about the BVED which are easily verified, will be used. Let $\lambda=\lambda_1 + \lambda_2 + \lambda_3$.

$$E_{R1}(X) = \lambda_1/\lambda^2$$

$$E_{R2}(X) = 1/(\lambda_1+\lambda_3) - (\lambda_1+\lambda_3)/\lambda^2$$

$$E_{R3}(X) = \lambda_3/\lambda^2$$

$$E(X) = 1/(\lambda_1+\lambda_3)$$

$$E_{R1}(Y) = 1/(\lambda_2+\lambda_3) - (\lambda_2+\lambda_3)/\lambda^2$$

$$E_{R2}(Y) = \lambda_2/\lambda^2$$

$$E_{R3}(Y) = \lambda_3/\lambda^2$$

$E(Y) = 1/(\lambda_2+\lambda_3)$

$E_{R1}(X^2) = 2\lambda_1/\lambda^3$

$E_{R2}(X^2) = 2/(\lambda_1+\lambda_3)^2 - 2(\lambda_1+\lambda_3)/\lambda^3$

$E_{R3}(X^2) = 2\lambda_3/\lambda^3$

$E(X^2) = 2/(\lambda_1+\lambda_3)^2$

$E_{R1}(Y^2) = 2/(\lambda_2+\lambda_3)^2 - 2(\lambda_2+\lambda_3)/\lambda^3$

$E_{R2}(Y^2) = 2\lambda_2/\lambda^3$

$E_{R3}(Y^2) = 2\lambda_3/\lambda^3$

$E(Y^2) = 2/(\lambda_2+\lambda_3)^2$

$E_{R1}(1) = \lambda_1/\lambda$

$E_{R2}(1) = \lambda_2/\lambda$

$E_{R3}(1) = \lambda_3/\lambda$

$E_{R1}(XY) = 2\lambda_1/\lambda^3 + \lambda_1/\lambda^2(\lambda_2+\lambda_3)$

$E_{R2}(XY) = 2\lambda_2/\lambda^3 + \lambda_2/\lambda^2(\lambda_1+\lambda_3)$

$E(XY) = \lambda_3/\lambda(\lambda_1+\lambda_3)(\lambda_2+\lambda_3) + 1/(\lambda_1+\lambda_3)(\lambda_2+\lambda_3)$

$E(\max(X,Y)) = 1/(\lambda_1+\lambda_3) + 1/(\lambda_2+\lambda_3) - 1/\lambda$

$E([\max(X,Y)]^2) = 2/(\lambda_1+\lambda_3)^2 + 2/(\lambda_2+\lambda_3)^2 - 2/\lambda^2$

$$E(S_1) = \underset{R1}{E}(1/\lambda_1 - X) + \underset{R2}{E}(1/(\lambda_1+\lambda_3) - X) + \underset{R3}{E}(-X)$$

$$= \underset{R1}{E}(1)/\lambda_1 + \underset{R2}{E}(1)/(\lambda_1+\lambda_3) - E(X)$$

$$= 1/\lambda + (\lambda_2-\lambda)/\lambda(\lambda_1+\lambda_3) = 1/\lambda - 1/\lambda = 0$$

$$E(S_2) = \underset{R1}{E}(1/(\lambda_2+\lambda_3) - Y) + \underset{R2}{E}(1/\lambda_2 - Y) + \underset{R3}{E}(-X)$$

$$= \underset{R1}{E}(1)/(\lambda_2+\lambda_3) + \underset{R2}{E}(1)/\lambda_2 - E(Y)$$

$$= \lambda_1/\lambda(\lambda_2+\lambda_3) + 1/\lambda - 1/(\lambda_2+\lambda_3) = 0$$

$$E(S_3) = \underset{R1}{E}(1/(\lambda_2+\lambda_3) - Y) + \underset{R2}{E}(1/(\lambda_1+\lambda_3) - X) + \underset{R3}{E}(1/\lambda_3 - X)$$

$$= \underset{R1}{E}(1)/(\lambda_2+\lambda_3) + \underset{R2}{E}(1)/(\lambda_1+\lambda_3) + \underset{R3}{E}(1)/\lambda_3 - E(\max(X,Y))$$

$$= \lambda_1/\lambda(\lambda_2+\lambda_3) + \lambda_2/\lambda(\lambda_1+\lambda_3) + 1/\lambda -$$

$$(1/(\lambda_1+\lambda_3) + 1/(\lambda_2+\lambda_3) - 1/\lambda) = 0$$

Since $E(S_i) = 0$ $(i = 1, 2, 3)$, it is verified that the BVED is regular with respect to its first partial-λ_i derivatives in Ω_3.

Again using the expressions for S_1, S_2, S_3 derived in Appendix A the following array of S_{ij}'s is obtained, where the three entries for each S_{ij} are the values of S_{ij} in regions R1, R2 and R3 from top to bottom. Since $S_{ij}=S_{ji}$, S_{21}, S_{31}, S_{32} are not listed.

$$S_{11} = \begin{cases} -1/\lambda_1^2 \\ -1/(\lambda_1+\lambda_3)^2 \\ 0 \end{cases} \qquad S_{12} = \begin{cases} 0 \\ 0 \\ 0 \end{cases} \qquad S_{13} = \begin{cases} 0 \\ -1/(\lambda_1+\lambda_3)^2 \\ 0 \end{cases}$$

$$S_{22} = \begin{cases} -1/(\lambda_2+\lambda_3)^2 \\ -1/\lambda_1^2 \\ 0 \end{cases} \qquad S_{23} = \begin{cases} -1/(\lambda_2+\lambda_3)^2 \\ 0 \\ 0 \end{cases} \qquad S_{33} = \begin{cases} -1/(\lambda_2+\lambda_3)^2 \\ -1/(\lambda_1+\lambda_3)^2 \\ -1/\lambda_3^2 \end{cases}$$

Let $D_{ij}=E(S_{ij})$ and D be the 3 x 3 matrix with entries D_{ij}. Then -D =

$$\lambda^{-1}\begin{pmatrix} 1/\lambda_1 + \lambda_2/(\lambda_1+\lambda_3)^2 & 0 & \lambda_2/(\lambda_1+\lambda_3)^2 \\ 0 & 1/\lambda_2 + \lambda_1/(\lambda_2+\lambda_3)^2 & \lambda_1/(\lambda_2+\lambda_3)^2 \\ \lambda_2/(\lambda_1+\lambda_3)^2 & \lambda_1/(\lambda_2+\lambda_3)^2 & 1/\lambda_3 + \lambda_1/(\lambda_2+\lambda_3)^2 + \lambda_2/(\lambda_1+\lambda_3)^2 \end{pmatrix}$$

Let $B_{ij}=E(S_iS_j)$ and B be the 3 x 3 matrix with entries B_{ij}. To show that the BVED is regular with respect to its second partial-λ_i derivatives in Ω_3 it must be verified that $B + D = 0_{3x3}$. The computation of the B_{ij} is routine but tedious. Only two terms B_{11} and B_{23} will be computed to illustrate the method. The other computations are similar and not interesting.

$$B_{11} = E(S_1^2) = \underset{R1}{E}(X - 1/\lambda_1)^2 + \underset{R2}{E}(X - 1/(\lambda_1+\lambda_3))^2 + \underset{R3}{E}(X^2)$$

$$= \underset{R1}{E}(X^2) - 2\underset{R1}{E}(X)/\lambda_1 + \underset{R1}{E}(1)/\lambda_1^2 + \underset{R2}{E}(X^2) - 2\underset{R2}{E}(X)/(\lambda_1+\lambda_3) + \underset{R2}{E}(1)/(\lambda_1+\lambda_3)^2 + \underset{R3}{E}(X^2)$$

$$= E(X^2) - 2\underset{R1}{E}(X)/\lambda_1 + \underset{R1}{E}(1)/\lambda_1^2 - 2\underset{R2}{E}(X)/(\lambda_1+\lambda_3) + \underset{R2}{E}(1)/(\lambda_1+\lambda_3)^2$$

$$= 2/(\lambda_1+\lambda_3)^2 - 2/\lambda^2 + 1/\lambda_1\lambda - 2[1/(\lambda_1+\lambda_3) - (\lambda_1+\lambda_3)/\lambda^2]/(\lambda_1+\lambda_3) + \lambda_2/\lambda(\lambda_1+\lambda_3)^2$$

$$= 1/\lambda_1\lambda + \lambda_2/\lambda(\lambda_1+\lambda_3)^2 = [1/\lambda_1 + \lambda_2/(\lambda_1+\lambda_3)^2]/\lambda = -D_{11}$$

$$B_{23} = E(S_2S_3) = \underset{R1}{E}(1/(\lambda_2+\lambda_3)-Y)^2 + \underset{R2}{E}(1/\lambda_2-Y)(1/(\lambda_1+\lambda_3)-X) + \underset{R3}{E}(X)(X-1/\lambda_3)$$

$$= \underset{R1}{E}(1)/(\lambda_2+\lambda_3)^2 - 2\underset{R1}{E}(Y)/(\lambda_2+\lambda_3) + \underset{R1}{E}(Y^2) + \underset{R2}{E}(1)/\lambda_2(\lambda_1+\lambda_3)$$

$$- \underset{R2}{E}(Y)/(\lambda_1+\lambda_3) - \underset{R2}{E}(X)/\lambda_2 + \underset{R2}{E}(XY) + \underset{R3}{E}(X^2) - \underset{R3}{E}(X)/\lambda_3$$

$$B_{23} = \lambda_1/\lambda(\lambda_2+\lambda_3)^2 - 2[1/(\lambda_2+\lambda_3) - (\lambda_2+\lambda_3)/\lambda^2]/(\lambda_2+\lambda_3)$$

$$+ [2/(\lambda_2+\lambda_3)^2 - 2(\lambda_2+\lambda_3)/\lambda^3] + 1/\lambda(\lambda_1+\lambda_3) - \lambda_2/(\lambda_1+\lambda_3)\lambda^2$$

$$- [1/(\lambda_1+\lambda_3) - (\lambda_1+\lambda_3)/\lambda^2]/\lambda_2 + [2\lambda_2/\lambda^3 + \lambda_2/\lambda^2(\lambda_1+\lambda_3)]$$

$$+ 2\lambda_3/\lambda^3 - 1/\lambda^2$$

$$B_{23} = \lambda_1/\lambda(\lambda_2+\lambda_3)^2 - 2/(\lambda_2+\lambda_3)^2 + 2/\lambda^2 + 2/(\lambda_2+\lambda_3)^2 - 2(\lambda_2+\lambda_3)/\lambda^3$$

$$+ 1/\lambda(\lambda_1+\lambda_3) - \lambda_2/(\lambda_1+\lambda_3)\lambda^2 - 1/\lambda_2(\lambda_1+\lambda_3) + (\lambda_1+\lambda_3)/\lambda_2\lambda^2$$

$$+ 2\lambda_2/\lambda^3 + \lambda_2/\lambda^2(\lambda_1+\lambda_3) + 2\lambda_3/\lambda^3 - 1/\lambda^2$$

$$= \lambda_1/\lambda(\lambda_2+\lambda_3)^2 + 1/\lambda^2 + 1/\lambda(\lambda_1+\lambda_3) - 1/\lambda_2(\lambda_1+\lambda_3) + (\lambda_1+\lambda_3)/\lambda_2\lambda^2$$

$$= \lambda_1/\lambda(\lambda_2+\lambda_3)^2 + [1/\lambda^2 + (\lambda_1+\lambda_3)/\lambda_2\lambda^2] - (1/\lambda_2 - 1/\lambda)/(\lambda_1+\lambda_3)$$

$$= \lambda_1/\lambda(\lambda_2+\lambda_3)^2 + [\lambda_2/\lambda + (\lambda_1+\lambda_3)/\lambda]/\lambda_2\lambda - 1/\lambda_2\lambda$$

But $\lambda_2/\lambda + (\lambda_1+\lambda_3)/\lambda = 1$ so $B_{23} = \lambda_1/\lambda(\lambda_2+\lambda_3)^2$ and thus $B_{23} + D_{23} = 0$.

In a similar fashion it can be verified that $B_{ij} + D_{ij} = 0$ for all other $i, j = 1, 2, 3$.

Hence the BVED is also regular with respect to its second partial-λ_i derivatives in Ω_3.

It should be noted that the matrix $B = -D$ multiplied by the sample size is the information matrix $I_n(\lambda_1, \lambda_2, \lambda_3)$ for the distribution based on a sample of size n. This has been given by Arnold [2] previously but no verification of the regularity conditions was indicated.

APPENDIX C

SOLUTION OF THE MAXIMUM LIKELIHOOD EQUATIONS

An iterative method to solve the maximum likelihood equations when $n_1n_2n_3>0$ is given here. Equations (3) of Chapter IV have the following form. Let a, b, c, d, e, f denote positive constants.

$$a/x + b/(x+z) = d$$

$$b/y + a/(y+z) = e$$

$$c/z + a/(y+z) + b/(x+z) = f$$

Let (x_o, y_o, z_o) be an initial guess for the solution. Expand each of the terms in x, y and z through linear terms about (x_o, y_o, z_o). Let Δx, Δy, Δz denote $x-x_o$, $y-y_o$, $z-z_o$, respectively. Then solve the resulting linear system below for Δx, Δy, Δz and let $(x_1, y_1, z_1) = (x_o+\Delta x, y_o+\Delta y, z_o+\Delta z)$.

$$a[1/x_o - \Delta x/x_o^2] + b[1/(x_o+z_o) - \Delta x/x_o^2 - \Delta z/z_o^2] = d$$

$$b[1/y_o - \Delta y/y_o^2] + a[1/(y_o+z_o) - \Delta y/y_o^2 - \Delta z/z_o^2] = e$$

$$c[1/z_o - \Delta z/z_o^2] + a[1/(y_o+z_o) - \Delta y/y_o^2 - \Delta z/z_o^2]$$
$$+ b[1/(x_o+z_o) - \Delta x/x_o^2 - \Delta z/z_o^2] = f$$

If the initial guess (x_o, y_o, z_o) is sufficiently close to (x, y, z) then this method of iteration converges satisfactorily. (This is essentially Newton's method which converges quadratically). Taking the initial guess for $\hat{\lambda}_1$, $\hat{\lambda}_2$, $\hat{\lambda}_3$ to be the estimates suggested by Arnold [2] or the estimates based on moments described in Chapter V, it has been found that convergence occurred in few iterations.

APPENDIX D

DETAILS OF EXPANSION FOR PROP ESTIMATES

The equations (5) referred to on page 35 of Chapter V are repeated here.

$$\hat{\lambda}_1 = f_1(a,b,c) = (1/a - c/b)/(1+c)$$
$$\hat{\lambda}_2 = f_2(a,b,c) = (1/b - c/a)/(1+c)$$
$$\hat{\lambda}_3 = f_3(a,b,c) = c(1/a + 1/b)/(1+c)$$

The method involves the following computations for each $i = 1, 2, 3$

$$\lambda_i = f_i(a,b,c) \simeq f_i(a_o,b_o,c_o) + f_{i1}(a_o,b_o,c_o)(a-a_o) + f_{i2}(a_o,b_o,c_o)(b-b_o) + f_{i3}(a_o,b_o,c_o)(c-c_o)$$

where f_{ij} denotes the partial of f_i with respect to its j^{th} argument. Note that $f_i(a_o,b_o,c_o) = \lambda_i$ for $i = 1, 2, 3$.

In the matrix notation suggested in Chapter V then

$$\hat{L} - L \simeq \left(f_{ij}(a_o,b_o,c_o)\right)_{3x3} \begin{pmatrix} a-a_o \\ b-b_o \\ c-c_o \end{pmatrix} = B\left(\sum Z_j/n - E(Z)\right)$$

So the ij^{th} entry in matrix B is the partial of f_i with respect to its j^{th} argument evaluated at (a_o,b_o,c_o).

$$f_{11} = -1/a_o^2(1+c_o) \qquad f_{12} = c_o/b_o^2(1+c_o)$$

$$f_{13} = -(1/a_o + 1/b_o)/(1+c_o)^2 \qquad f_{21} = c_o/a_o^2(1+c_o)$$

$$f_{22} = -1/b_o^2(1+c_o) \qquad f_{23} = -(1/a_o + 1/b_o)/(1+c_o)^2$$

$f_{31} = -c_o/a_o^2(1+c_o)$ $\qquad$ $f_{32} = -c_o/b_o^2(1+c_o)$

$f_{33} = (1/a_o + 1/b_o)/(1+c_o)^2$

Upon substituting for a_o, b_o, c_o, the values $1/(\lambda_1+\lambda_3)$, $1/(\lambda_2+\lambda_3)$, λ_3/λ, respectively, the terms given in the B matrix in Chapter IV may be verified.

Now expanding f_i through quadratic terms in $(a-a_o)$, $(b-b_o)$, $(c-c_o)$ then subtracting $f_i(a_o, b_o, c_o)$ from both sides and taking expected values, one obtains approximate biases for the PROP estimates.

$$E(\hat{\lambda}_i-\lambda_i) \approx \sum_{j=1}^{3}\sum_{k=1}^{3} f_{ijk}(a_o,b_o,c_o)\mathrm{Cov}(A_j,A_k)/2 \qquad (i = 1,\ 2,\ 3)$$

where f_{ijk} denotes the second partial of function f_i with respect to its j^{th} and k^{th} argument and A_1=a, A_2=b, A_3=c. $\mathrm{Cov}(A_j, A_k)$ is given by the (j,k) entry in Cov(Z) divided by n found in Chapter V. The details will be given for i=1 only, the others are similar. All partials below are evaluated at (a_o, b_o, c_o).

$f_{111} = 2/a_o^3(1+c_o)$ $\qquad$ $f_{112} = 0$

$f_{113} = 1/a_o^2(1+c_o)^2$ $\qquad$ $f_{121} = f_{112}$

$f_{122} = -2c_o/b_o^3(1+c_o)$ $\qquad$ $f_{123} = 1/b_o^2(1+c_o)^2$

$f_{131} = f_{113}$ $\qquad$ $f_{132} = f_{123}$

$f_{133} = 2(1/a_o + 1/b_o)/(1+c_o)^3$

Substituting the values $1/(\lambda_1+\lambda_3)$, $1/(\lambda_2+\lambda_3)$, $\lambda_3/\lambda = \rho$ for a_o, b_o, c_o into the above expressions and simplifying the algebra it is found

$$E(\hat{\lambda}_1-\lambda_1) \simeq [2(\lambda_1+\lambda_3)/(1+\rho) - 2(\lambda_1+\lambda_3)\lambda_2\lambda_3/\lambda^2(1+\rho)^2$$

$$- 2\rho(\lambda_2+\lambda_3)/(1+\rho) - 2(\lambda_2+\lambda_3)\lambda_1\lambda_3/\lambda^2(1+\rho)^2$$

$$+ 2(\lambda+\lambda_3)\lambda_3(\lambda_1+\lambda_2)/\lambda^2(1+\rho)^3]/2n$$

After cancelling a factor of 2, the first and third terms inside the brackets may be combined to give λ_1 while the remainder of the terms in the brackets simplify to $\rho(\lambda_1{}^2+\lambda_2{}^2)/\lambda(1+\rho)^2$.

Hence $E(\hat{\lambda}_1-\lambda_1)$ is approximately $[\lambda_1 + \rho(\lambda_1{}^2+\lambda_2{}^2)/\lambda(1+\rho)^2]/n$. Or $E(\hat{\lambda}_1) \simeq \lambda_1 + \lambda_1/n + \rho(\lambda_1{}^2+\lambda_2{}^2)/\lambda(1+\rho)^2n$. That is, the expected value of $\hat{\lambda}_1$ is of the order $\lambda_1(1+1/n)$. An approximate correction for bias then would be to consider $n\hat{\lambda}_1/(n+1)$.

APPENDIX E

SIMULATION OF SAMPLES FROM THE BIVARIATE EXPONENTIAL DISTRIBUTION

If $F(x,y)$ is the joint cumulative distribution function of the random variables X and Y then a random sample of size one (X, Y) from this distribution may be obtained in the following manner. First consider the marginal distribution of X, $F_1(x)$. It is well known that $F_1(X)$ has a uniform distribution on [0,1]. Let U_1 and U_2 be a random sample of size two from the uniform distribution. Then $X = F_1^{-1}(U_1)$ provided that the inverse function F_1^{-1} can be obtained in closed form. Then consider the conditional distribution $F(Y|X)$. For the fixed X the random variable $F(Y|X)$ also has a uniform distribution. Let $G(Y) = F(Y|X)$. Then $Y = G^{-1}(U_2)$ provided that the inverse function G^{-1} can be obtained in closed form. In this way a random sample of size 1 (X, Y) may be obtained from $F(x, y)$. (The roles of X and Y may be interchanged; that is, obtain a value for Y first and then a value for X.) To obtain a random sample of size n, a random sample of size 2n is required from the uniform distribution. If the distribution is the bivariate exponential distribution then it was seen in Chapter III that this procedure could be followed.

The special property pointed out in [1] that $(X, Y) \sim BVE(\lambda_1, \lambda_2, \lambda_3)$ if and only if there exist three independent random variables V_1, V_2, and V_3 each with exponential distributions with parameters λ_1, λ_2, λ_3 respectively, such that $X = \min(V_1, V_3)$ and $Y = \min(V_2, V_3)$, permits an easier method to simulate a sample of size one (X, Y) from $BVE(\lambda_1, \lambda_2, \lambda_3)$. Let U_1, U_2, U_3 be a random sample of size three from the uniform distribution. $F(V_i) = 1 - \exp\{-\lambda_i V_i\}$ and $F(V_i)$ has a uniform distribution for $i = 1, 2, 3$. So $V_i = -\ln(U_i)/\lambda_i (\lambda_i > 0)$ for $i = 1, 2, 3$. Then $X = \min(U_1, U_3)$ and $Y = \min(U_2, U_3)$. If $\lambda_3 = 0$ then U_3 is not required and $X = -\ln(U_1)/\lambda_1$, $Y = -\ln(U_2)/\lambda_2$. To obtain a sample of size n from $BVE(\lambda_1, \lambda_2, \lambda_3)$ with $\lambda_1, \lambda_2, \lambda_3 > 0$ a random sample of size 3n is required from the uniform distribution in [0, 1]. The above procedure was used in this thesis to generate samples from the bivariate exponential distribution.

The U_i's were generated by an I.B.M. supplied subroutine available on the U.M.R. computer. The fact that these U_i's are representative of a uniform distribution has been well documented in [18] and [19]. The number of samples from the BVED used for each sample size n found in tables in this thesis is given below.

Sample Size	10	20	40	100	200
Number of Samples	8000	4000	2000	1000	500

By studying control cases these numbers were obtained as reasonable numbers of samples for each sample size indicated.

For example, for n=20 when 4000 samples were generated from $BVE(\gamma_1, \gamma_2, \rho)$ with $\gamma_1 = \gamma_2 = 1$, $\rho = .1$, the means and variances of the simulated values of certain statistics are compared below with their true values. For the definitions of these statistics see Chapters III and IV.

The percent error in the simulated means compared to the exact values does not exceed .5% and the percent error in the simulated variances compared to the exact values does not exceed 2.5%. While it is difficult to make a precise statement about the simulated values

Statistic	Simulated Mean	Simulated Variance	True Mean	True Variance
n_1	9.010	4.959	9.000	4.950
n_2	8.994	5.064	9.000	4.950
n_3	1.995	1.823	2.000	1.800
AX	19.97	19.96	20.00	20.00
AY	20.03	20.02	20.00	20.00
AZ	28.98	26.10	29.00	25.85

of other statistics whose distributions may be unknown, it is felt that the errors are of the same order of magnitude. Clearly generating larger numbers of samples would be desirable but limitations of available computer time made this unfeasible.

REFERENCES

1. Marshall, A.W. and Olkin, I. (1966), A Multivariate Exponential Distribution. Journal of the American Statistical Association, 61, 30-44.

2. Arnold, B.C. (1968), Parameter Estimation for a Multivariate Exponential Distribution. Journal of the American Statistical Association, 63, 848-852.

3. Thompson, W.A. (1969), Applied Probability. Holt, Rinehart, Winston, New York, New York, 125-131.

4. Esary, J.D. and Proschan, F. (1970), A Reliability Bound for Systems of Maintained, Interdependent Components. Journal of the American Statistical Association, 65, 329-338.

5. Esary, J.D., Proschan, F., Walkup, D.W. (1967), Association of Random Variables, with Applications. Annals of Mathematical Statistics, 38, 1466-1474.

6. Harris, R. (1970), A Multivariate Definition for Increasing Hazard Rate Distribution Functions. Annals of Mathematical Statistics, 41, 713-717.

7. Maik, R. (1969), Applications of Order Statistics to the Multivariate Exponential Distribution, Abstract 6. Annals of Mathematical Statistics, 40, 725.

8. Maik, R. (1970), Estimating the Parameters of a Multivariate Exponential Using Order Statistics, Abstract 70T-13. Annals of Mathematical Statistics, 41, 754.

9. Kibble, W.F. (1941), A Two-Variate Gamma Type Distribution. Sankhyā 5, 137-150.

10. Wilks, S.S. (1962), Mathematical Statistics. John Wiley and Sons, Inc., New York, New York.

11. Tables of the Bivariate Normal Distribution Function and Related Functions, (1959). U.S. Department of Commerce, National Bureau of Standards, Washington D.C.

12. Harter, H.L. (1964), New Tables of the Incomplete Gamma-Function Ratio and of Percentage Points of the Chi-Square and Beta Distributions. Aerospace Research Laboratories, U.S.A.F., Washington D.C.

13. Cramer, H. (1946), Mathematical Methods of Statistics. Princeton University Press, Princeton, New Jersey.

14. Loeve, M.M. (1955), Probability Theory, Van Nostrand, New York, New York.

15. Lehmann, E.L. (1959), Testing Statistical Hypotheses. John Wiley and Sons, Inc., New York, New York.

16. Bain, L.J. (1967), Tolerance Regions for a Joint Exponential Distribution. I.E.E.E. Transactions on Reliability R-16, 111-113.

17. Kendall, M.C., Stuart A. (1961), The Advanced Theory of Statistics. Vol. II, Griffin and Co., Ltd., London, England.

18. Haas, G.N. (1969), Statistical Inferences for the Cauchy Distribution Based on Maximum Likelihood Estimators. Thesis, University of Missouri at Rolla, 102 p. (with sixteen figures, 33 tables).

19. Dumonceaux, R.H. (1969), Statistical Inferences for Location and Scale Parameter Distributions. Thesis, University of Missouri at Rolla, 83 p. (with 4 figures, 55 tables).

20. Barlow, R.E., Proschan, F. (1965), Mathematical Theory of Reliability. John Wiley and Sons, Inc., New York, New York.

21. Epstein, B. (1958), The Exponential Distribution and its Role in Life Testing. Industrial Quality Control, 15, 4-9.

22. Esary, J.D., Marshall, A.W. (1964), System Structure and the Existence of a System Life. Technometrics 6, 459-462.

23. Hogg, R.V., Craig, A.T. (1965), Introduction to Mathematical Statistics, Second Edition. The MacMillan Company, New York, New York.

24. Linnik, J.V. (1968), Statistical Problems with Nuisance Parameters. American Mathematical Society, Providence, Rhode Island.

25. Lloyd, D.K. and Lipow, M. (1962), Reliability: Management, Methods, and Mathematics. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

VITA

Bruce Mohr Bemis was born on June 30, 1936, in San Francisco, California. He received his primary and secondary education in San Francisco graduating from Washington High School in 1954. His undergraduate work was done at Westminster College in Salt Lake City, Utah where he graduated magna cum laude in 1958 receiving a Bachelor of Arts degree with a major in mathematics. He served as a graduate assistant at the University of Utah from 1958 to 1960 while pursuing his graduate studies there and received the Master of Science in mathematics in August 1960. For the next two years he was employed at Hercules, Incorporated in Magna, Utah as an Evaluation Engineer.

In the fall of 1962 he took a position as Instructor of Mathematics at Westminster College, Salt Lake City, Utah. In 1964 he was promoted to Assistant Professor and Chairman of the Mathematics Department.

From 1968 to 1970 he was enrolled in the Graduate School of the University of Missouri - Rolla and was a Science Faculty Fellow during the academic year 1968-69 and held a N.A.S.A. Traineeship the second year. He is a member of the American Statistical Association and the National Council of Teachers of Mathematics.

In 1959 he married Joy Marie Hammer and they have three children, Ronald, Kathryn and Sheryl.